PUBLIC HEALTH INFORMATICS — BIOSURVEILLANCE AND OPERATIONS STRATEGIES

A Dissertation
Presented to
The Academic Faculty

By

Yifan Liu

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
H. Milton Stewart School of Industrial and Systems Engineering

Georgia Institute of Technology

May 2018

PUBLIC HEALTH INFORMATICS — BIOSURVEILLANCE AND OPERATIONS STRATEGIES

Approved by:

Dr. Eva K. Lee, Advisor
H. Milton Stewart School of Industrial and Systems Engineering
Georgia Institute of Technology

Dr. Andy Sun
H. Milton Stewart School of Industrial and Systems Engineering
Georgia Institute of Technology

Dr. David Goldsman
H. Milton Stewart School of Industrial and Systems Engineering
Georgia Institute of Technology

Dr. D. J. Wu
Ernest Scheller Jr. College of Business
Georgia Institute of Technology

Dr. Yajun Mei
H. Milton Stewart School of Industrial and Systems Engineering
Georgia Institute of Technology

Date Approved: March 28, 2018

ACKNOWLEDGEMENTS

First, I would like to express my gratitude to my advisor, Dr. Eva K. Lee, for her supervision, guidance, and financial support. The accomplishment of this dissertation, as well as my personal and intellectual growth throughout the doctoral study, would be impossible without her.

My sincere appreciation also goes to Dr. David Goldsman, Dr. Yajun Mei, Dr. Andy Sun, and Dr. D. J. Wu for serving on my dissertation committee. I am grateful for the helpful discussion and valuable feedback. I would also acknowledge Ferdinand H. Pietz, Dr. Samuel B. Graitcer, and Dr. Lisa M. Koonin from Centers for Disease Control and Prevention for sponsoring my research and reviewing my work.

I gratefully acknowledge the funding sources that made my doctoral study possible. The work was partially supported by Centers for Disease Control and Prevention and the National Science Foundation.

I would like to thank the faculty members who taught me the professional knowledge and the staff members who provided me with administrative and technical support, especially Amanda Ford and Mark Danielson. My special thanks go to my fellow students who studied with me and banged heads with me: Rui Gao, Weijun Xie, Jikai Zou, Helin Zhu, Xiaowei Yue, Yang Cao, Joshua Hale, and Satya Sarvani Malladi. I wish to express my indebtedness to all my friends who helped and supported me throughout this journey. To Yuno Cheng and Willie Fishman, thank you for the encouragement in my difficult times.

Lastly, I would like to express my deepest appreciation to my parents for their unconditional support and love. I have been blessed to have parents like you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Public health emergencies continue to pose serious threats to human lives and well-being of the world population. The 2014 Ebola virus outbreak in West Africa and the 2016 Zika virus outbreak in South and North America underscores the challenges post by some of these highly infectious diseases even as healthcare has been revolutionized over the past century with medical science and technology. Infectious diseases are one of the most common yet most serious types of public health crisis due to several reasons: high infectivity, difficulties in prevention, labor-intensive intervention, diagnosis, and treatment. This is amplified by the mobility of human population enabled by transportation technologies. To react efficiently and effectively to waves of infectious disease outbreaks, decision-makers and public health officers must quickly analyze the current situation and predict the potential trend of an outbreak, evaluate multiple countermeasures and strategies, and implement an intervention plan that optimizes utilization of the available (limited) resource while achieving the best containment results. Additionally, these decisions need to be made in real time as any delay in this process may result in severe outcomes, including more infections, economic loss, and even lives loss.

In this dissertation, we advance infectious diseases models with applications in medical countermeasure operations and biosurveillance. We first propose a general-purpose modeling framework for infectious disease outbreaks and discuss how it can be used to facilitate decision making to achieve best containment results. We demonstrate how this modeling technique can be computerized and used to support public health operations.

In the first chapter, we describe a general-purpose modeling framework for infectious disease. By expanding and abstracting the traditional compartmental models widely used for disease propagation dynamics, our modeling framework is highly generic and can be

viewed as a metamodel for compartmental models. We focus on the models for contact-based diseases and vector-borne diseases and apply it to two real-life scenarios: 1) the Zika virus outbreaks in Brazil and Puerto Rico, and 2) the avian influenza outbreaks in Egypt and the United States. We discuss how this generic modeling framework can be tailored and parameterized to suit the properties of different diseases at different regions and investigate the impact of different intervention strategies.

In the second chapter, we equip the disease modeling framework with an optimization engine to determine the optimal resource allocation strategy during an outbreak. We investigate the impact of point-of-dispensing (PODs) for rapid medical countermeasures dispensing during an outbreak. We analyze its usage for a regional smallpox outbreak in the state of Georgia and investigate vaccination strategies for maximum health protection.

In the third chapter, we explain the design and functionalities of an enterprise software package, the RealOpt suite. Designed at Georgia Institute of Technology in collaboration with the Centers for Disease Control and Prevention since 2005, RealOpt is a modularized system. This chapter focuses on one aspect of RealOpt: an expanded module known as RealOpt-Regional. We discuss the design structure of the general RealOpt framework and its ease-of-usage by public health personnel for emergency preparedness and actual responses. We then demonstrate usage in multiple aspects of public health emergency response, including facility location optimization, resource allocation optimization, inventory and transportation management, and cost-effectiveness analysis of strategies.

CHAPTER 1
DISEASE PROPAGATION MODELS FOR BIOSURVEILLANCE

In this chapter, we propose a general-purpose modeling framework for infectious disease propagation with applications in public health practice, biosurveillance, and medical research. This modeling framework is an extension and generalization of the compartmental models and can be used to monitor the spread of pandemic outbreaks, evaluate containment strategies, and help decision-makers understand the biological properties of different diseases. We discuss two applications of this framework in modeling and containing Zika virus outbreaks and avian influenza outbreaks.

1.1 Literature Review

Infectious diseases continue to be one of the major causes of mortalities in many countries as of the 21st century. They are the third leading cause of death in the United States and the leading cause worldwide [29]. In addition, the infectious agents such as bacteria and viruses continuously evolve and create new challenges for human beings [164]. The recent Ebola virus outbreak in West Africa in 2014 and the Zika virus outbreak in South America and The Caribbean in 2016 indicate that combating emerging infectious diseases is still one of the most important tasks for emergency response.

Mathematical models have been shown to be important techniques in understanding the underlying mechanisms of the spread and control of infectious diseases by utilizing the available information and transforming it into the knowledge of the disease, then providing strategies and guidelines to prevent the disease from becoming a global pandemic. Coupled with computer simulation, mathematical models are powerful tools to provide valuable insights into disease outbreaks, understand the transmission characteristics, test hypothesis,

and assess and evaluate containment strategies. Mathematical epidemiology models have been studied extensively for many years, dating back to the smallpox model first formulated by Daniel Bernoulli in the 18th century [28].

The foundations of mathematical models for infectious diseases were developed in the late 19th and early 20th centuries [34]. These studies on mathematical models can be briefly separated into two categories: deterministic and stochastic epidemiology models.

1.1.1 Deterministic Epidemiology Models

Deterministic models have long been applied to study the infectious disease epidemiology. Hamer developed a discrete time model to understand the recurrence of measles in 1906, in which he assumed that the number of new infections depends on the product of the density of susceptible and infectious populations [107]. This idea sets the foundation for future compartmental models. In 1911, Ross developed a system of differential equations using a host-vector structure for the control of malaria [187]. Other deterministic models were also developed for multiple purposes in the early 20th century [16], until Kermack and McKendrick published a series of papers in the 1920s and 1930s, developing the theory of SIR models and other compartmental models [123–125, 158].

Kermack–McKendrick Model

The original form of Kermack–McKendrick model is a compartmental differential equation model consists of susceptible, infectious, and recovered populations with coefficients relative to the distribution of age in the population. Kermack and McKendrick also studied the threshold conditions for this system for the occurrence of an epidemic outbreak. Many mathematical models for pandemics were developed afterwards based on their framework. For example, Capasso and Serio generalized the deterministic compartmental model with an interaction term and characterized the occurrence of infections with a nonlinear bounded

function [41]. Instead of unlimited contacts between individuals at a uniform rate, Diekmann et al. introduced a population contact structure and studied the effect of repeated contacts within a group of "acquaintances" [65]. Clancy and Piunovskiy studied the model that replaces the infection rate function in the original Kermack–McKendrick model with a dynamic infection rate adjusted by the total population in the system and studied the optimal isolation policy to intervene under different conditions [51].

Other extensions and generalizations based on the Kermack–McKendrick theory are still topics focused by recent researchers. Li and Zou studied the compartmental model in which the infectious disease has a fixed latent period and formulated the SIR structure for a population living in two patches [149]. Pathak et al. replaced the constant infection rate with an asymptotically homogeneous transmission function and derived the stability condition of the model [177]. Xu developed a diffusive Kermack–McKendrick epidemic model with a latent period and determined conditions of the existence of traveling waves solutions of the system [216]. Inaba extended the compartmental model to recognize individual heterogeneity, expanding the definition of compartments to genetic, psychological, or behavioral characteristics of population [111]. Disraelly et al. extended the methodology of Human Response Injury Profile (HRIP) which uses time-based progression to determine casualty and fatality estimations from infections. Based on the compartmental model, they introduced an injury profile sub-model based on the severity of the symptoms to describe the progression of illness at a detailed granularity [69]. Refer to Allen [8], Daley and Gani [60], Hethcote [110], and Breda et al. [36] for complete reviews of the compartmental epidemiology models proposed based on the work of Kermack and McKendrick.

Population-Dependent Models

The original Kermack–McKendrick model has age-dependent recovery and transmission rates. While the theory of the simple SIR model treats these parameters as constants,

population-dependent compartmental models are also well studied. Brauer considered models with nonlinear population dynamics with permanent removal and studied the stability of endemic equilibrium the model [35]. Pugliese studied a compartmental model with no recovery and density-dependent mortality and incidence rate and examined the effect of vaccination [180]. Zhou and Hethcote analyzed an SIS type and considered various demographic structures including recruitment-death, logistic, decay and growth [218]. Roberts described the spread of fatal infectious diseases with a system of nonlinear integral equations regulated by population density-dependent constraints and compared this model with the ODE-based compartmental models [186]. Kuperman and Abramson modeled different structures of populations and their interactions within small world networks and studied the spread of infections within the network [131]. For a complete review of population-dependent dynamic disease models, see Anderson and May [11].

Models with Passive Immunity and Vaccination

With the flexibility of compartmental models, additional stages representing passive immunity, vaccination, hospitalization or quarantine can be introduced to the system to help determine the best practice to achieve early containment. Since these stages are not natural to the propagation of infectious diseases, special models are needed to analyze their effectiveness and efficiency. Schuette and Hethcote used an age-structured model to study and evaluate the effects of varicella vaccination programs [188]. Li and Ma studied an SIS epidemic model with vaccination, temporary immunity, and dynamical population size, and identified three threshold parameters that determine the equilibrium of the system [148]. Arino et al. formulated a disease model that incorporate the movement of individuals over spatial scales and discussed the influence of quarantine in the form of travel restriction [14]. Arino et al. formulated compartmental models for influenza with control by vaccination and antiviral treatment, and compared the analytical result with that of stochastic simulations [13]. Liu et al. developed an SVIR structure to evaluate continuous vaccina-

tion strategy and pulse vaccination strategy for vaccine doses that need to be taken several times [152].

Besides modeling the effect of interventions on a large scale, the effects of internal disease propagation within hospitals, clinics, and point-of-dispensing sites (PODs) are also studied. Liao et al. employed an SIR structure and a dynamic transmission model to quantify the risk associated with indoor airborne infection [150]. Lee et al. analyzed the propagation of highly infectious diseases within dispensing sites with large-scale simulation [135]. The effect of prioritizing vaccine administration to the high-risk population when the supply of vaccine is limited was studied by Lee et al. based on a combination of compartmental models and a queueing model [143].

1.1.2 Stochastic Epidemiology Models

Stochastic models emphasize the randomness in the spread of infectious diseases and model it as a stochastic process, or add stochastic components in the deterministic models. They are preferred compared to deterministic models when analyses are possible since it is more natural to depict the spread of disease as stochastic, and stochastic models are useful in parameter estimations and hypothesis testings [12]. In addition, simple deterministic models may not be useful in understanding underlying principles of the spread of diseases in some cases [113]. For complete reviews of stochastic epidemiology models, see Andersson and Britton [12] and Britton [37].

Disease Models with Stochastic Process

The stochastic process approach was first developed by Russian physician P. D. En'ko. He fitted a discrete chain binomial model for the spread of measles in St. Petersburg in 1889 [66]. This model was further developed and referred as Reed-Frost model in the 1920's named after Lowell Reed and Wade Hampton Frost from Johns Hopkins University

[1]. The Reed-Frost model describes the spread of infections in generations, and each infectious individuals in generation t infects susceptible individuals independently with probability p. Some asymptotic properties of the Reed-Frost model were studied by Sellke [190] and Ball and O'Neill [19]. A similar approach is the Greenwood model proposed in 1931 [103], which models the number of susceptible population S_t at generation t as a binomial random variable that depends on both S_{t-1} and a probability p: $S_t \sim \mathcal{B}(S_{t-1}, p)$.

Markov chain is also widely used to model epidemiology. Under the closed population assumption, there are finite number of states and each state represents a possible distribution of the susceptible and infectious populations. However, it was difficult to achieve an analytical solution to this system [22]. Kendall provided an approximation to the Markovian compartmental model and studied it analytically [122]. Other analytical results of simpler Markov models in both discrete time and continuous time were developed and studied [9, 165, 198]. Cooper and Lipsitch applied hidden Markov models to time series to model hospital-acquired infections caused by transmissible pathogens [55]. Rao et al. introduced state transitions in the Markov processes to model spatial interactions between entities under an agent-based simulation framework [182].

Disease Models with Stochastic Differential Equations

Stochastic differential equations (SDE) were widely used to capture the randomness in the spread of infectious diseases. One family of SDE models was developed based on their deterministic counterparts by adding stochastic components. For example, based on the Kermack–McKendrick framework, models with nonlinear disease propagation trend [49, 83] and population with recruitment and deaths [116] were later developed. Allen constructed a stochastic differential equation approximation to the continuous-time Markov chain model in an SIS setting [7]. Stochastic diffusion model was also developed and was shown as equivalent to a system of ordinary differential equations with time-variant

coefficients [196].

Another method to introduce stochasticity to the deterministic models is parameter perturbation. The application of parameter perturbation in SIS models was discussed by Dalal et al. [59] and Gray et al. [102]. Analytical properties such as conditions for stability and asymptotic behaviors of the system on more complex compartmental models were also studied. Chen and Li [48] and Tornatore et al. [197] investigated the stability of random SIR models; Ding et al. studied the asymptotic behavior of the stochastic model specifically for AIDS [67]; the properties of the stochastic SIRS model was studied by Lu [153].

Disease Models with Simulation or Network Structure

With the availability of computational power in recent decades, more complex stochastic models for infectious diseases were studied. Simulation is widely used in studying the complicated underlying mechanism in disease spread. Focks et al. developed a pair of stochastic simulation models to study the epidemiology of dengue fever in the urban environment [89]. Smith et al. used individual-based stochastic simulation to predict the development and evaluate interventions of malaria outbreaks [194]. Nakamura et al. studied the disease spreading with an agent-based simulation model analytically and enabled the evaluation of stationary states in Markov processes [167]. The application of multi-agent simulation in spatial epidemiology was explored by Ostfeld et al. [173] and Dion et al. [68].

Monte Carlo simulation was also used to investigate epidemic events by treating it as a non-equilibrium dynamic process, or fitting parameters for complex models. Hass et al. described the disease spread in an SIR setting using global and local variables in Monte Carlo simulation with different settings of square lattices [104]. The similar approach was adapted by Karsten et al. to investigate the swine fever epidemics and their control [119].

Aiello and da Silva used dynamical Monte Carlo methods to simulate Markov processes based on a generalized SIRS model, and the performance of this approach agreed with the 4th order Runge-Kutta Method for the deterministic model [4]. Markov Chain Monte Carlo model was used in estimating epidemiological parameters of H1N1 influenza [200] and deducting model parameters for stochastic compartmental models from incomplete observations [96].

The network structure plays an important role in modeling the spread of diseases, as it has better representations of the spatial distribution of infections. Bansal et al. used a network perspective to relax the homogeneous-mixing assumption which is commonly used in mathematical epidemiology models and evaluated several methods to incorporate contact heterogeneity [20]. Theoretical analysis of epidemic spreading on heterogeneous networks was studied, and the potential extensions towards coevolving, coupled, and time-varying networks were discussed by Pastor-Satorras et al. [175]. May and Lloyd demonstrated properties of infection processes on scale-free networks, and discussed the conditions of achieving threshold behavior in such models [156]. Keeling and Eames described a series of methods that allow the approximation and computation of mixing networks [121].

1.2 Modeling Framework for Contact-Based Diseases

In this section, we discuss the general-purpose modeling framework for contact-based diseases. We first categorize the disease stages or compartments according to their roles and properties in the spread of infectious diseases; then we define the essential variables for the general-purpose model based on the categorization of stages; finally, we propose the deterministic modeling framework for contact-based diseases.

1.2.1 Categorization of Disease Stages

The general-purpose modeling framework for infectious diseases is developed based on the compartmental models. Compartments, or disease stages, are the statuses that individuals can fall into during the spread of infectious diseases. Assume that the human population is homogeneous. Let Φ denote the collection of all possible stages. For example, for an SIR model, $\Phi = \{S, I, R\}$. Then any stage in Φ can be categorized in two different ways:

* Passive/Active stages. Individuals in a passive stage will not change their statuses spontaneously. For instance, the susceptible stage is passive, as susceptible individuals remain susceptible unless they make contact with the infectious population and become infected. Let Φ_P denote the collection of passive stages. On the contrary, individuals in an active stage will change their statuses spontaneously. An example of active stages is infectious: infectious individuals will either recover or decease given sufficiently long time. Let Φ_A denote the collection of active stages. Obviously, a stage is either passive or active. Therefore, $\Phi_P \cup \Phi_A = \Phi$. This method categorization determines if a mean dwelling time is well defined for a stage.

* Vulnerable/Contagious/Unaffected stages. Individuals in a vulnerable stage can be infected through contacting the infectious ones. Let Φ_V denote the collection of all vulnerable stages. Individuals in a contagious stage have the ability to infect vulnerable entities. We use Φ_C to denote all the contagious stages. The third category is unaffected stages, in which individuals are neither vulnerable nor contagious, and they are not in the system of infection. Denote them as Φ_U. For example, after the initial infection, entities will enter the exposed stage. They are not vulnerable, as they have already been infected; but they are not infectious neither since the density of viruses or bacteria has not reached the threshold to infect others. Another example of the unaffected stage is the recovered stage in an SIR model with immunity assumed. All natural stages without intervention will fall into one of these categories, thus

$$\Phi_V \cup \Phi_C \cup \Phi_U = \Phi.$$

Therefore, for each disease stage, it can be categorized using these two methods, and these two ways of categorization capture the major characteristics of a disease stage. Take the 4-stage SEIR (susceptible-exposed-infectious-recovered) model as an example. Figure 1.1 shows the transition diagram of the model and the categorization of each stage.

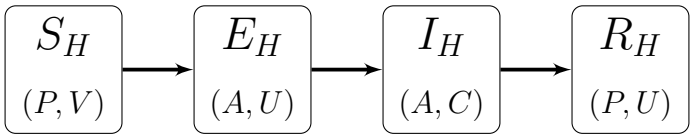


Figure 1.1: Demonstration of stage categorization with an SEIR model.

In the SEIR setup, the passive/active stage categorization is

$$\Phi_P = \{S, R\}, \Phi_A = \{E, I\}$$

and the vulnerable/contagious/unaffected stage categorization is

$$\Phi_V = \{S\}, \Phi_C = \{I\}, \Phi_U = \{E, R\}$$

1.2.2 Definition of Model Components

Contact-based diseases are transmitted via physical or indirect contacts between humans. The main interest of compartmental model is to understand how the population associated with each disease stage changes with respect to time t. Assume the total population in the system is N at the beginning of an outbreak. Let $\phi(t)$ denote the number of individuals in the corresponding compartment at time t, for all $\phi \in \Phi$. Then if we ignore the natural birth and death of population, we have $\sum_{\phi\in\Phi} \phi(t) = N$ for any given t.

Define vector $\boldsymbol{y} = \langle \phi(t) \rangle|_{\phi\in\Phi}$ as the population in each compartment at time t. Simi-

larly, the first-order derivative of the population in each stage in terms of time t is be written as

$$\boldsymbol{y}' = \frac{\mathrm{d}\boldsymbol{y}}{\mathrm{d}t} = \left\langle \frac{\mathrm{d}\phi(t)}{\mathrm{d}t} \right\rangle \Bigg|_{\phi \in \Phi}$$

For all active stages, there is a well-defined average transition rate for each of them. This transition rate represents how long an individual will stay in an active stage before transiting to another stage. Therefore, for all $\phi \in \Phi_A$, let μ_ϕ denote the mean transition rate for stage ϕ; for $\phi \in \Phi_P$, define $\mu_\phi = 0$ for completeness. Let $\boldsymbol{\mu} = \langle \mu_\phi \rangle|_{\phi \in \Phi}$ be the mean transition rate for each stage.

To understand the destination of the disease transitions, define a *disease transition matrix* $\boldsymbol{D} = \langle d_{ab} \rangle|_{a \in \Phi, b \in \Phi}$ such that d_{ab} represents the disease transition probability from stage b to stage a. The column sum of matrix $\boldsymbol{D}$ is either 0 or 1, where stages with column sum equal to 0 are absorbing stages. Since the transition structure of an infectious disease is determined by its own biological property, the disease transition matrix $\boldsymbol{D}$ can be viewed as a constant parameter. Notice that new infections are also considered as disease transition. For example, in the SEIR model discussed in Figure 1.1, the disease transition matrix is

$$\boldsymbol{D} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Let β_ϕ denote the effective baseline infection rate or contact rate adjusted by the total population in the system for all vulnerable stages $\phi \in \Phi_V$. Again, define $\beta_\phi = 0$ for stages $\phi \in \phi_C \cup \phi_U$ for completeness. Use $\boldsymbol{\beta} = \langle \beta_\phi \rangle|_{\phi \in \Phi}$ to denote the baseline contact rate for every stage. Unlike the disease transition matrix $\boldsymbol{D}$, the value of vector $\boldsymbol{\beta}$ may change with respect to time t due to the change in the distribution of the population. An intuitive

example is that people normally do not have contacts with the deceased population, thus the deceased population needs to be excluded when calculating the population-adjusted contact rate $\boldsymbol{\beta}$.

In the previous section, we defined the relationship between vulnerable and contagious stages: individuals in contagious stages can infect those in vulnerable stages via direct or indirect contacts, depending on the biological property of the infectious disease. Here we represent this relationship mathematically. For simplicity, we define a function $\mathcal{S}(\phi) : \Phi \to 2^{\Phi}$ such that $\mathcal{S}(\phi)$ gives the set of successor stages of ϕ in the disease transition diagram. Since the transition structure can be fully characterized by the disease transition matrix $\boldsymbol{D}$, the function $\mathcal{S}(\phi)$ can be expressed equivalently as

$$\mathcal{S}(\phi) = \{\psi \in \Phi : d_{\psi\phi} > 0\} \tag{1.1}$$

Similarly, we can define a function $\mathcal{P}(\phi) : \Phi \to 2^{\Phi}$ such that $\mathcal{P}(\phi)$ gives the set of predecessor stages of ϕ. By using the definition of matrix $\boldsymbol{D}$, the function $\mathcal{P}(\phi)$ can be expressed as

$$\mathcal{P}(\phi) = \{\psi \in \Phi : d_{\phi\psi} > 0\} \tag{1.2}$$

With function $\mathcal{S}$ properly defined, we can derive the *disease contagious matrix* $\boldsymbol{C} = \langle c_{ab} \rangle|_{a\in\Phi, b\in\Phi}$ to represent the relationship between contagious and vulnerable stages. In particular, $c_{ab} = -1$ if and only if $a \in \Phi_V$ and $b \in \Phi_C$; $c_{kb} = 1$ for all $k \in \mathcal{S}(a)$ if and only if $a \in \Phi_V$ and $b \in \Phi_C$. For example, the disease contagious matrix for the SEIR model is

$$\boldsymbol{C} = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Table 1.1 summarizes all the model components discussed in this section. These components will be used to establish the general-purpose modeling framework.

Table 1.1: Definition of model components for modeling disease propagations.

Symbol	Definition
$\boldsymbol{y} = \langle \phi(t) \rangle|_{\phi \in \Phi}$	Number of entities in each compartment at time t.
$\boldsymbol{y}' = \langle \phi'(t) \rangle|_{\phi \in \Phi}$	The derivative of populations in each stage with respect to time t.
$\boldsymbol{\mu} = \langle \mu_\phi \rangle|_{\phi \in \Phi}$	Mean transition rate for active stages. Define $\mu_\phi = 0$ for passive stages for completeness.
$\boldsymbol{\beta} = \langle \beta_\phi \rangle|_{\phi \in \Phi}$	Baseline infection rate for vulnerable stages adjusted by total population. Define $\beta_\phi = 0$ for all $\phi \notin \Phi_V$ for completeness.
$\boldsymbol{D} = \langle d_{ab} \rangle|_{a \in \Phi, b \in \Phi}$	Disease transition matrix: d_{ab} denotes the transition probability from stage b to stage a; fully characterizes the transition diagram.
$\boldsymbol{C} = \langle c_{ab} \rangle|_{a \in \Phi, b \in \Phi}$	Disease contagious matrix: if $a \in \Phi_V$ and $b \in \Phi_C$, then $c_{ab} = -1$ and $c_{kb} = 1$ for all successor stages k of a.

1.2.3 Modeling Framework for Contact-Based Diseases

With the model components defined in Table 1.1, we now derive the general-purpose modeling framework for the spread of contact-based infectious diseases. In particular, we would like to derive the expression for $\boldsymbol{y}'$ using these modeling components, which gives us a convenient way to compute the population changes in each compartment within a time unit.

The changes in population can be decomposed into two parts: the changes due to the natural progression of infectious diseases, and the changes due to new infections. For the first part, the change of population in stage ϕ due to the natural development of itself can be written as $-\mu_\phi \phi(t)$. It can be rewritten more compactly in the matrix form: $-\text{diag}(\boldsymbol{\mu})\boldsymbol{y}$.

This only considers the populations flowing out of the active stages. Since disease transition matrix $\boldsymbol{D}$ captures the transition structure, it can be used to describe the destinations of these population changes: the change of population in each stage due to the inflow caused by the natural progression of active stages can be written as $\boldsymbol{D}\text{diag}(\boldsymbol{\mu})\boldsymbol{y}$. By combining the population changes in these two parts, the change of populations in each stage due to the natural development of infectious disease can be written as:

$$\boldsymbol{y}'_{\text{development}} = (\boldsymbol{D} - \boldsymbol{I})\text{diag}(\boldsymbol{\mu})\boldsymbol{y} \tag{1.3}$$

The discussion above concludes the population change due to the natural development of infectious diseases in the first part. For the second part, the cause of population changes is new infections due to direct or indirect contacts between vulnerable and contagious individuals. The generation of new infections is determined by the populations of both vulnerable and contagious stages. Consider a vulnerable individual in a vulnerable stage $\phi \in \Phi_V$ who makes b_ϕ contacts with others on average in a unit time. Assume the total population that can be reached by this individual at time t is $N(t)$, and $C(t)$ out of these $N(t)$ individuals are contagious. Therefore, the average size of contagious population contacted by this individual in a unit time is $b_\phi C(t)/N(t)$. Since the total size of population in the vulnerable stage ϕ is $\phi(t)$, the total new infections introduced to the system in a unit time from vulnerable stage ϕ is $b_\phi \phi(t) C(t)/N(t)$. Since the baseline infection rate $\boldsymbol{\beta}$ defined in Table 1.1 is already adjusted by the total population $N(t)$, the size of new infections from vulnerable stage ϕ can be written as $\beta_\phi \phi(t) C(t)$.

Therefore, for a vulnerable stage $\phi \in \Phi_V$, the changes of population due to new infections is $-\beta_\phi \phi(t) \sum_{\psi \in \Phi_C} \psi(t)$. The infection structure between vulnerable and contagious stages has already been captured by disease contagious matrix $\boldsymbol{C}$, thus this change of population can be rewritten compactly in matrix form as $\text{diag}(\text{diag}(\boldsymbol{\beta})\boldsymbol{y})\boldsymbol{C}\boldsymbol{y}$. Similarly, this

equation only considers the outflow of population from vulnerable stages due to new infections. To capture the destinations of these flows, the disease transition matrix $\boldsymbol{D}$ is used. The inflow in each stage due to new infection is then

$$\text{diag}(\boldsymbol{D}\text{diag}(\boldsymbol{\beta})\boldsymbol{y})\boldsymbol{C}\boldsymbol{y} \tag{1.4}$$

Notice that this equation can be justified using Equation 1.1 and Equation 1.2: for a vulnerable stage $v \in \Phi_V$ that can be infected by individuals in stage $u \in \Phi_C$, $c_{\psi v} = 1$ for all $\psi \in \mathcal{S}(v)$. Then it is trivial that for each $\psi \in \mathcal{S}(v)$,

$$\boldsymbol{e}_\psi^\intercal \boldsymbol{C}\boldsymbol{y} = \sum_{\omega \in \Phi_C} \omega(t) \tag{1.5}$$

where $\boldsymbol{e}_\psi$ is the standard unit vector with the element corresponding to stage ψ being 1.

On the other hand, for each $\psi \in \mathcal{S}(v)$,

$$\boldsymbol{e}_\psi^\intercal \boldsymbol{D}\text{diag}(\boldsymbol{\beta})\boldsymbol{y} = \sum_{\phi \in \mathcal{P}(\psi) \cap \Phi_V} d_{\psi\phi}\beta_\phi\phi(t) \tag{1.6}$$

Combining Equation 1.5 and Equation 1.6, for a stage ψ that is a successor of any vulnerable stage,

$$\boldsymbol{e}_\psi^\intercal \text{diag}(\boldsymbol{D}\text{diag}(\boldsymbol{\beta})\boldsymbol{y})\boldsymbol{C}\boldsymbol{y} = \sum_{\phi \in \mathcal{P}(\psi) \cap \Phi_V} d_{\psi\phi}\beta_\phi\phi(t) \sum_{\omega \in \Phi_C} \omega(t)$$

Notice that the column sum of disease transition matrix $\boldsymbol{D}$ is 1 for any vulnerable stage. Adding up this value for all stages with predecessors of vulnerable stages,

$$\begin{aligned}\sum_{\psi:\mathcal{P}(\psi)\cap\Phi_V \neq \varnothing} \boldsymbol{e}_\psi^\intercal \text{diag}(\boldsymbol{D}\text{diag}(\boldsymbol{\beta})\boldsymbol{y})\boldsymbol{C}\boldsymbol{y} &= \sum_{\psi:\mathcal{P}(\psi)\cap\Phi_V \neq \varnothing} \sum_{\phi \in \mathcal{P}(\psi)\cap\Phi_V} d_{\psi\phi}\beta_\phi\phi(t) \sum_{\omega \in \Phi_C} \omega(t) \\ &= \sum_{\phi \in \Phi_V} \beta_\phi\phi(t) \sum_{\omega \in \Phi_C} \omega(t)\end{aligned}$$

which is the opposite of the summation of outflows in vulnerable stages. Therefore, the validity of Equation 1.4 is justified.

By combining the population changes in these two parts, the change of population in each stage due to the introduction of new infections in the system can be written as:

$$\boldsymbol{y}'_{\text{infection}} = \text{diag}((\boldsymbol{D} + \boldsymbol{I})\text{diag}(\boldsymbol{\beta})\boldsymbol{y})\boldsymbol{C}\boldsymbol{y} \tag{1.7}$$

We have derived the equation to model the population change due to the natural progression of diseases in Equation 1.3 and the equation to describe the population change due to new infections in Equation 1.7. By adding up these two equations, we obtain a compact expression of $\boldsymbol{y}'$, which concludes our general-purpose modeling framework for contact-based diseases:

$$\boldsymbol{y}' = \boldsymbol{y}'_{\text{development}} + \boldsymbol{y}'_{\text{infection}} = (\boldsymbol{D} - \boldsymbol{I})\text{diag}(\boldsymbol{\mu})\boldsymbol{y} + \text{diag}((\boldsymbol{D} + \boldsymbol{I})\text{diag}(\boldsymbol{\beta})\boldsymbol{y})\boldsymbol{C}\boldsymbol{y} \tag{1.8}$$

Notice that this equation does not include natural born and death of the population. An additional term $\lambda(\boldsymbol{y})$ can be added to Equation 1.8 to describe the change of population in each stage due to the natural born and death. Other factors that may cause the population change can also be introduced to the current framework.

1.2.4 Model Stochasticity with Langevin Dynamics

Although the proposed modeling framework for contact-based infectious diseases described in Equation 1.8 is deterministic, stochasticity can be introduced into this framework to model the uncertainty in the spread of diseases. Langevin dynamics, first proposed by French physicist Paul Langevin in 1908 [146], is a system of stochastic differential equations with the stochastic terms accounting for omitted degrees of freedom. It was originally used to model the dynamics of molecular systems in the field of statistic physics.

This modeling technique is widely used to describe the evolution of the probabilities associated to system dynamics with large population in each compartment in the study of chemistry and epidemiology [52, 53, 97].

The idea of Langevin dynamics is to divide the variables in the system into two groups: macroscopic variables change slowly and are responsible for the changes in mean values of the system; microscopic variable changes faster with respect to time compared to the macroscopic variables, and contribute to the stochastic part of the model. A generic form of Langevin dynamics is

$$\boldsymbol{F}(t) = \boldsymbol{y}'(t) + \boldsymbol{\eta}(t)$$

where $\boldsymbol{y}'(t)$ is the derivative of the macroscopic variable $\boldsymbol{y}$ with respect to time t, and $\boldsymbol{\eta}(t)$ is a Gaussian noise. Usually, the magnitude of the Gaussian noise is proportional to the square root of its macroscopic counterpart [94, 97]. Therefore, for a given stage ϕ in the modeling framework for infectious diseases, the Langevin dynamics can be written as

$$\phi(t) = \phi'(t) + \sqrt{\phi'(t)}\eta \tag{1.9}$$

where $\eta \sim \mathcal{N}(0,1)$. Denote $H(\boldsymbol{A}) = \boldsymbol{A}^{\circ\frac{1}{2}}$ as the Hadamard square root of nonnegative matrix $\boldsymbol{A}$, i.e., the (i,j) entry of $H(\boldsymbol{A})$ is $\sqrt{a_{ij}}$. Then the general-purpose model for contact-based infectious disease with Langevin dynamic can be derived directly from Equation 1.8 and Equation 1.9:

$$\begin{aligned}\boldsymbol{y}' =&(\boldsymbol{D}-\boldsymbol{I})\text{diag}(\boldsymbol{\mu})\boldsymbol{y} + \text{diag}((\boldsymbol{D}+\boldsymbol{I})\text{diag}(\boldsymbol{\beta})\boldsymbol{y})\boldsymbol{C}\boldsymbol{y}\\ &-(\boldsymbol{D}-\boldsymbol{I})\text{diag}(\boldsymbol{\eta}_D)H(\text{diag}(\boldsymbol{\mu})\boldsymbol{y}) - H(\text{diag}((\boldsymbol{D}+\boldsymbol{I})\text{diag}(\boldsymbol{\beta})\boldsymbol{y}))\boldsymbol{C}\text{diag}(\boldsymbol{\eta}_C)H(\boldsymbol{y})\end{aligned}$$

where $\boldsymbol{\eta}_D$ and $\boldsymbol{\eta}_C$ are two statistically independent Gaussian random vectors with mean $\boldsymbol{0}$ and variance-covariance matrix $\boldsymbol{I}$, representing the stochastic parts in the population

change in each stage corresponding to disease progression and new infections. Notice that both matrices $\text{diag}(\boldsymbol{\mu})\boldsymbol{y}$ and $\text{diag}((\boldsymbol{D}+\boldsymbol{I})\text{diag}(\boldsymbol{\beta})\boldsymbol{y})$ are nonnegative, thus the Hadamard square root in the equation above is well-defined.

1.3 Modeling Framework for Vector-Borne Diseases

Vector-borne diseases are transmitted via different types of vectors and do not usually transmit directly between humans. Examples of vector-borne diseases include dengue, West Nile virus, yellow fever, and Zika virus. Corresponding compartmental models have been developed and studied [61, 79–82, 84]. In this section, we develop a general-purpose modeling framework for vector-borne infectious diseases and discuss some properties and potential extensions and generalizations of this model.

1.3.1 Stages and Model Parameters for Human and Vector Populations

For simplicity, assume there are only two groups of populations in the system: human and vector. Susceptible humans get infected through contact with infectious vectors, and infectious human can infect susceptible vectors through contact as well. Assume an SEIR structure for the human population, and an SID structure for the vector population, where "D" stands for deceased. Figure 1.2 demonstrate a disease transition diagram for this setting, where subscript H stands for the human stages, and subscript V stands for the vector stages. The black solid arrows refer to the natural progression of diseases, and the red dashed arrows refer to contacts between humans and vectors that cause infections.

Let Φ^H denote the collection of all disease stages in the human population, and let Φ^V denote all stages in the vector population. Let $\boldsymbol{y}_H$ and $\boldsymbol{y}_V$ denote the population of each compartment in both human and vector groups. The categorization of disease stages discussed in Section 1.2.1 still work with the separation of human and vector stages. Since the symptoms and outcomes of infection are different between humans and vectors, all model

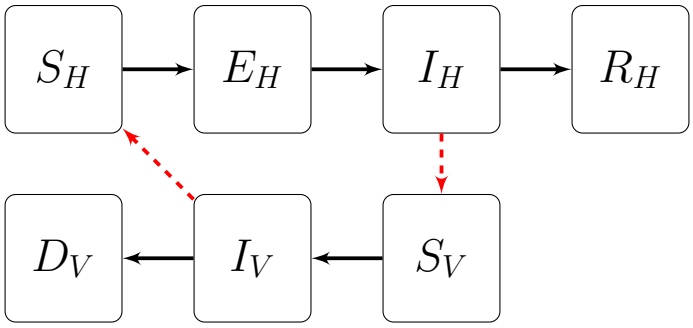


Figure 1.2: Demonstration of stage transitions for a vector-borne disease.

parameters need to be distinguished by population groups. Let $\boldsymbol{\mu}_H$ and $\boldsymbol{\mu}_V$ denote the mean transition rate for active stages in human population and vector population respectively, and let $\mu_\phi = 0$ for passive stages in both human and vector population. Similarly, the disease transition matrices for human and vector population are denoted as $\boldsymbol{D}_H$ and $\boldsymbol{D}_V$.

Let $\boldsymbol{\beta}_{V\to H}$ and $\boldsymbol{\beta}_{H\to V}$ denote the baseline infection rates for human and vector population due to the inter-species contacts adjusted by the total human population in the system. Though the baseline contact rates between the two populations are the same, the probabilities of getting infected may differ. Consider the scenario in which the overall contact rate between two groups is b, and the probabilities of getting infectious from these contacts in each disease stage is $\boldsymbol{p}_H$ for human and $\boldsymbol{p}_V$ for vector, then $\boldsymbol{\beta}_{V\to H} = b\boldsymbol{p}_H, \boldsymbol{\beta}_{H\to V} = b\boldsymbol{p}_V$. Besides the inter-species contacts, intra-species contacts may also introduce new infections to the system, depending on the biological properties of the disease. Therefore, let $\boldsymbol{\beta}_{H\to H}$ and $\boldsymbol{\beta}_{V\to V}$ be the baseline infection rate due to the contacts within the same population group. These two parameters have the same interpretation as the baseline infection rate $\boldsymbol{\beta}$ in the model for contact-based diseases.

Since the infection structure involves the interaction of two populations, there should be two different disease contagious matrices, one for each direction of the infection. Let $\boldsymbol{C}_{V\to H} = \langle c_{ab}\rangle|_{a\in\Phi^H, b\in\Phi^V}$ denote the disease contagious matrix for vectors infecting hu-

mans, where $c_{ij} = -1$ if $i \in \Phi_V^H, j \in \Phi_C^V$, and $c_{kj} = 1$ for all $k \in \mathcal{S}(i)$ if $i \in \Phi_V^H, j \in \Phi_C^V$. Similarly, let $\boldsymbol{C}_{H\to V} = \langle c_{ab}\rangle|_{a\in\Phi^V, b\in\Phi^H}$ be the disease contagious matrix for humans infecting vectors, where $c_{ij} = -1$ if $i \in \Phi_V^V, j \in \Phi_C^H$, and $c_{kj} = 1$ for all $k \in \mathcal{S}(i)$ if $i \in \Phi_V^V, j \in \Phi_C^H$. If infections within human or vector population group are possible, the intra-group disease contagious matrices $\boldsymbol{C}_{H\to H}$ and $\boldsymbol{C}_{V\to V}$ can be defined in the same way as in the model for contact-based diseases.

1.3.2 Modeling Framework for Vector-Borne Diseases

Similar to the modeling framework for contact-based diseases, the model for the vector-borne disease can be separated into two parts: the population change in stages due to the natural disease progression, and the population change due to new infections. Since the natural progression of diseases does not involve the interaction between humans and vectors, this part can be written independently for both human and vector population:

$$\boldsymbol{y}'_{H,\text{development}} = (\boldsymbol{D}_H - \boldsymbol{I})\text{diag}(\boldsymbol{\mu}_H)\boldsymbol{y}_H, \boldsymbol{y}'_{V,\text{development}} = (\boldsymbol{D}_V - \boldsymbol{I})\text{diag}(\boldsymbol{\mu}_V)\boldsymbol{y}_V$$

The change in populations due to new infections can be further split into two parts: the intra-species infections and the inter-species infections. The approach to model intra-species infections is the same as the model for contact-based diseases, as this change does not involve the interactions between species:

$$\boldsymbol{y}'_{H,\text{intra-infection}} = \text{diag}((\boldsymbol{D}_H + \boldsymbol{I})\text{diag}(\boldsymbol{\beta}_{H\to H})\boldsymbol{y}_H)\boldsymbol{C}_{H\to H}\boldsymbol{y}_H$$
$$\boldsymbol{y}'_{V,\text{intra-infection}} = \text{diag}((\boldsymbol{D}_V + \boldsymbol{I})\text{diag}(\boldsymbol{\beta}_{V\to V})\boldsymbol{y}_V)\boldsymbol{C}_{V\to V}\boldsymbol{y}_V$$

To model the inter-species infections, assume the total human and vector populations at time t are $N_H(t)$ and $N_V(t)$, respectively. Consider a vulnerable stage ϕ for the human population. Since the baseline contact rate of vectors infecting humans adjusted by the human population in stage ϕ is $\beta_{V\to H,\phi}$, the number of infectious contacts a human make

with the vector population in a unit time is $\beta_{V\to H,\phi}N_V(t)$. Among the $N_V(t)$ vectors, $C_V(t)$ of them are contagious. Therefore, the exposure rate to infections for a single individual in stage ϕ is $\beta_{V\to H,\phi}C_V(t)$.

Therefore, for a vulnerable stage ϕ in the human population, the change rate due to new infection is $-\beta_{V\to H,\phi}\phi(t)\sum_{\psi\in\Phi_C^V}\psi(t)$. Written more compactly with the disease contagious matrix $\boldsymbol{C}_{V\to H}$, the reductions of population in all human stages due to new infections can be written as $\text{diag}(\text{diag}(\boldsymbol{\beta}_{V\to H})\boldsymbol{y}_H)\boldsymbol{C}_{V\to H}\boldsymbol{y}_V$. Again, the destinations of the outflows due to new infections are captured in the disease transition matrix for human $\boldsymbol{D}_H$, thus the inflows for all human stages can be written as $\text{diag}(\boldsymbol{D}_H\text{diag}(\boldsymbol{\beta}_{V\to H})\boldsymbol{y}_H)\boldsymbol{C}_{V\to H}\boldsymbol{y}_V$, following the same reasoning in the model for contact-based diseases. Therefore, the change in populations in the human group due to inter-species infections is

$$\boldsymbol{y}'_{H,\text{inter-infection}} = \text{diag}((\boldsymbol{D}_H + \boldsymbol{I})\text{diag}(\boldsymbol{\beta}_{V\to H})\boldsymbol{y}_H)\boldsymbol{C}_{V\to H}\boldsymbol{y}_V$$

Similarly, for a vulnerable stage ϕ in the vector population, the baseline contact rate of humans infecting vectors in this stage adjusted by the human population is $\beta_{H\to V,\phi}$ per unit time. Assume the total contagious human population at time t is $C_H(t)$. Therefore, the exposure rate to infections for vectors in stage ϕ is $\beta_{H\to V,\phi}C_H(t)$. Then for the same vulnerable stage ϕ in the vector population, the change rate due to new infection is $-\beta_{H\to V,\phi}\phi(t)\sum_{\psi\in\Phi_C^H}\psi(t)$. Using the disease contagious matrix of human infecting vectors $\boldsymbol{C}_{H\to V}$, the reductions of populations in all vector stages due to new infections is $\text{diag}(\text{diag}(\boldsymbol{\beta}_{H\to V})\boldsymbol{y}_V)\boldsymbol{C}_{H\to V}\boldsymbol{y}_H$. With $\boldsymbol{D}_V$ capturing the information of the destinations of the outflows from stages due to new infections, the inflows for all vector stages is $\text{diag}(\boldsymbol{D}_V\text{diag}(\boldsymbol{\beta}_{H\to V})\boldsymbol{y}_V)\boldsymbol{C}_{H\to V}\boldsymbol{y}_H$. Summing up these two parts, the changes of population in the vector group due to inter-species infections are:

$$\boldsymbol{y}'_{V,\text{inter-infection}} = \text{diag}((\boldsymbol{D}_V + \boldsymbol{I})\text{diag}(\boldsymbol{\beta}_{H\to V})\boldsymbol{y}_V)\boldsymbol{C}_{H\to V}\boldsymbol{y}_H$$

By adding up the changes in population due to disease progression, intra-species infections and inter-species infections, the general-purpose modeling framework for vector-borne diseases can be written as:

$$\begin{aligned}\boldsymbol{y}'_H = &(\boldsymbol{D}_H - \boldsymbol{I})\text{diag}(\boldsymbol{\mu}_H)\boldsymbol{y}_H + \text{diag}((\boldsymbol{D}_H + \boldsymbol{I})\text{diag}(\boldsymbol{\beta}_{H\to H})\boldsymbol{y}_H)\boldsymbol{C}_{H\to H}\boldsymbol{y}_H \\ &+ \text{diag}((\boldsymbol{D}_H + \boldsymbol{I})\text{diag}(\boldsymbol{\beta}_{V\to H})\boldsymbol{y}_H)\boldsymbol{C}_{V\to H}\boldsymbol{y}_V \qquad (1.10a)\\ \boldsymbol{y}'_V = &(\boldsymbol{D}_V - \boldsymbol{I})\text{diag}(\boldsymbol{\mu}_V)\boldsymbol{y}_V + \text{diag}((\boldsymbol{D}_V + \boldsymbol{I})\text{diag}(\boldsymbol{\beta}_{V\to V})\boldsymbol{y}_V)\boldsymbol{C}_{V\to V}\boldsymbol{y}_V \\ &+ \text{diag}((\boldsymbol{D}_V + \boldsymbol{I})\text{diag}(\boldsymbol{\beta}_{H\to V})\boldsymbol{y}_V)\boldsymbol{C}_{H\to V}\boldsymbol{y}_H \qquad (1.10b)\end{aligned}$$

This modeling framework for vector-borne diseases can be further expanded to include multiple species of vectors by introducing proper model parameters and interaction terms between each combination of human and vector species. Additionally, the effect of alternative hosts for vector-borne diseases other than humans can also be introduced by adding compartments corresponding to the alternative hosts population or use their populations to adjust the baseline infection rate.

1.3.3 Modeling Dynamic Vector Population

In the previous models, the natural birth and death process of the human population was not discussed in details, except for using a generic function $\lambda(\boldsymbol{y}_H)$ to represent the natural birth and death process. Since the length of an infectious disease outbreak is usually short compared to the longevity of humans, the changes in human population due to reasons other than infections can be safely ignored without drastically altering the results. However, since the life cycles of common vectors (mosquitoes, fleas, etc.) are usually much shorter than that of humans, their life cycles can play an essential role in determining the development of disease outbreaks. Therefore, the definition of disease stages in the vector population Φ^V needs to be expanded to capture their life cycles.

Take mosquitoes as an example. There are four different stages of the entire life cycle of mosquitoes: eggs, larvae, pupae, and adults. Since the adult population has already been considered in our previous modeling framework, three additional stages representing eggs, larvae, and pupae should be introduced to Φ^V. Mosquitoes in these stages do not contact human directly, thus are not involved in the disease propagation, and we may assume they belong to the unaffected stages Φ_U^V. On the other hand, although entities in these stages will actively change their status and enter the next stage, they should not be categorized to the active stages Φ_A^V, since the spontaneous change of stage is due to natural progression instead of disease propagation. For more general modeling purposes, the belonging of infectious stages defined in previous sections needs to be determined according to the biological properties of the vector being studied.

For stage ϕ belonging to additional stages representing the dynamic vector population, if ϕ is active in terms of natural progression, there is a mean transition rate to other stages τ_ϕ associated to it. Assume $\boldsymbol{\tau}_V = \langle \tau_{ab} \rangle|_{\phi \in \Phi^V}$ denote the mean transition rate for each vector stages, while $\tau_\phi = 0$ if ϕ is not part of the stages representing natural population progressions. Define a *natural development matrix* $\boldsymbol{R}_V = \langle r_{ab} \rangle|_{a \in \Phi^V, b \in \Phi^V}$ such that $r_{ab} = 1$ if transition happens from stage b to stage a due to natural development, and $r_{ab} = 0$ otherwise. Again, the column sum of $\boldsymbol{R}_V$ should be either 1 or 0, while 0 represents an absorbing stage in the sense of natural progression without infections. Notice the difference between $\boldsymbol{R}_V$ and $\boldsymbol{D}_V$: $\boldsymbol{R}_V$ represents the population change due to natural development, which happens regardless of the existence of infections; $\boldsymbol{D}_V$ represents the population changes due to the development of infections and is well-defined if and only if an infectious disease is spreading. Therefore, following the same reasoning in deriving the equations for changes in populations due to the natural development of infections, the population changes

due to the dynamics of vector life cycle can be written as:

$$\boldsymbol{y}'_{V,\text{progression}} = (\boldsymbol{R}_V - \boldsymbol{I})\text{diag}(\boldsymbol{\tau}_V)\boldsymbol{y}_V$$

The second aspect in modeling the dynamic vector population is the reproduction process. Let $\boldsymbol{L}_V = \langle l_{ab} \rangle|_{a\in\Phi^V, b\in\Phi^V}$ denote the *reproduction indicator matrix*, where l_{ab} equals the reproduction rate if from stage b to stage a, i.e., vectors in stage b will give birth to new vector populations belonging to stage a at rate l_{ab}. Define $l_{ab} = 0$ if the reproduction process does not happen between stage b to stage a. Since the reproduction process only increases the vector population and does not impact the existing population, the population changes due to due to reproduction in the natural dynamics of vector population can be written as

$$\boldsymbol{y}'_{V,\text{reproduction}} = \boldsymbol{L}_V\boldsymbol{y}_V$$

Combining the two aspects of changes in vector population due to natural development and reproduction, the dynamic vector population can be modeled as

$$\boldsymbol{y}'_{V,\text{dynamic}} = ((\boldsymbol{R}_V - \boldsymbol{I})\text{diag}(\boldsymbol{\tau}_V) + \boldsymbol{L}_V)\boldsymbol{y}_V$$

and this term can be added to Equation 1.10b and provide a more detailed description of the vector population dynamics throughout an infectious disease outbreak.

1.4 Basic Reproduction Number

The basic reproduction number $\mathcal{R}_0$, defined as the average number of secondary infections caused by one infectious individual in a system consists of only susceptible population, is one of the major concepts in the research of epidemiology models. It is a measure of the potential for disease spread, as well as a threshold for the stability of a disease-free equilibrium of the ODE system [63]. If $\mathcal{R}_0 < 1$, the number of new infections introduced

to the system will fail to replace themselves, and the outbreak will start to contain; if $\mathcal{R}_0 > 1$, then the infectious population will increase and the disease will spread.

Castillo-Chavez et al. [45], Van den Driessche and Watmough [71, 72] and Diekmann et al. [64] used the approach of next generation matrix at equilibrium to define the basic reproduction number $\mathcal{R}_0$. Following this method, we derive the basic reproduction number for our general-purpose modeling framework in Equation 1.8. First, assume the system has a disease-free equilibrium $\boldsymbol{y}_\infty$ such that $\phi(\infty) = 0$ for all $\phi \in \Phi_A$, i.e., there is no population in active stages, thus no new infections will be introduced to the system spontaneously. Assume new infections are introduced to the system at this equilibrium, and the populations in all passive stages are treated as constants for simplicity. Therefore, we may only consider the population changes in active stages. Let $\boldsymbol{y}_A$ denote the partial population vector with only active stages of vector $\boldsymbol{y}$. Let $\boldsymbol{A} = \langle a_{\phi\psi} \rangle|_{\phi\in\Phi_A,\psi\in\Phi}$ be a $|\Phi_A| \times |\Phi|$ matrix such that $a_{\phi\phi} = 1$ if and only if $\phi \in \Phi_A$, while all other elements are 0. For any $|\Phi| \times |\Phi|$ matrix $\boldsymbol{B}$, $\boldsymbol{A}\boldsymbol{B}\boldsymbol{A}^\intercal$ gives a sub-matrix of $\boldsymbol{B}$ which consists only rows and columns in Φ_A. Following Equation 1.8, the population changes in active stages after introducing a new infection at the equilibrium $\boldsymbol{y}_\infty$ is

$$\boldsymbol{y}'_A = \boldsymbol{A}(\boldsymbol{D} - \boldsymbol{I})\text{diag}(\boldsymbol{\mu})\boldsymbol{A}^\intercal\boldsymbol{y}_A + \boldsymbol{A}\text{diag}((\boldsymbol{D} + \boldsymbol{I})\text{diag}(\boldsymbol{\beta})\boldsymbol{y}_\infty)\boldsymbol{C}\boldsymbol{A}^\intercal\boldsymbol{y}_A \tag{1.11}$$

Following the notation in [71] and [72], denote

$$\boldsymbol{V} = -\boldsymbol{A}(\boldsymbol{D} - \boldsymbol{I})\text{diag}(\boldsymbol{\mu})\boldsymbol{A}^\intercal, \boldsymbol{F} = \boldsymbol{A}\text{diag}((\boldsymbol{D} + \boldsymbol{I})\text{diag}(\boldsymbol{\beta})\boldsymbol{y}_\infty)\boldsymbol{C}\boldsymbol{A}^\intercal$$

then Equation 1.11 can now be written as $\boldsymbol{y}'_A = (\boldsymbol{F} - \boldsymbol{V})\boldsymbol{y}_A$. The number of infections produced by the new infectious individual is the product of the expected duration of this individual staying infectious and the rate of introducing new infections. The rate of introducing new infections is already captured in matrix $\boldsymbol{F}$. To calculate the expected duration

of staying infectious, consider the following system which only involves disease progression with a specified initial value $\boldsymbol{y}_0$ without new infections:

$$\boldsymbol{y}_A' = -\boldsymbol{V}\boldsymbol{y}_A, \boldsymbol{y}_A(0) = \boldsymbol{y}_0$$

The solution of this system is $h(t, \boldsymbol{y}_0) = e^{-\boldsymbol{V}t}\boldsymbol{y}_0$, where each component in this solution can be interpreted as the probability that the infectious individual represented by $\boldsymbol{y}_0$ is in the corresponding disease stage when introduced to the system at time $t = 0$. Therefore, the total number of new infections introduced by $\boldsymbol{y}_0$ is

$$\int_0^\infty \boldsymbol{F}h(t, \boldsymbol{y}_0)\mathrm{d}t = \boldsymbol{F}\int_0^\infty e^{-\boldsymbol{V}t}\mathrm{d}t\boldsymbol{y}_0 = \boldsymbol{F}\boldsymbol{V}^{-1}\boldsymbol{y}_0$$

The matrix $\boldsymbol{K} = \boldsymbol{F}\boldsymbol{V}^{-1}$ is called next generation matrix for the system at the disease-free equilibrium $\boldsymbol{y}_\infty$. The (ϕ, ψ) entry of matrix $\boldsymbol{K}$ is the expected number of new infections in stage ϕ produced by infectious individuals initially in stage ψ. Van den Driessche and Watmough showed that $\boldsymbol{K}$ has nonnegative eigenvalues, and the basic reproduction number of the system is defined as $\mathcal{R}_0 = \rho(\boldsymbol{K})$, where $\rho(\boldsymbol{K})$ denote the spectral radius of matrix $\boldsymbol{K}$, which is the maximum of the moduli of the eigenvalues of $\boldsymbol{K}$. The eigenvector $\boldsymbol{\omega}$ associated with $\mathcal{R}_0$ is also shown to be nonnegative [71, 72].

The idea behind this definition of $\mathcal{R}_0$ originates from the asymptotic behavior of the system. Assume after m generations after the initial introduction of infection individual $\boldsymbol{y}_0$, the magnitude of infectious population is $||\boldsymbol{K}^m\boldsymbol{y}_0||_1$, and the per-generation growth factor is $||\boldsymbol{K}^m||_1^{1/m}$. Let $m \to \infty$, then $\lim_{m\to\infty} ||\boldsymbol{K}^m||_1^{1/m} = \rho(\boldsymbol{K})$ by Gelfand's formula. Therefore, the average number of secondary infections, in the long run, is given by $\rho(\boldsymbol{K})$, which defines the basic reproduction number in the previous discussion. See Diekmann et al. [64] for more details. An interpretation of this definition of basic reproduction number is that if the distribution of the infectious individual introduced at the equilibrium follows

the eigenvector $\boldsymbol{\omega}$ of the matrix $\boldsymbol{K}$ that corresponds to eigenvalue $\mathcal{R}_0$, then the maximal number of typical secondary infections produced by the initial infection will be $\mathcal{R}_0$.

Notice that calculating the basic reproduction number with next generation matrix requires prior knowledge on the disease-free equilibrium $\boldsymbol{y}_\infty$. Such equilibrium may not be obtained analytically in many scenarios. In that case, numerical approaches are needed to compute the basic reproduction number. Take the SEIR model described in Figure 1.1 as an example and calculate its basic reproduction number. Since no closed-form disease-free equilibrium can be obtained for an arbitrary distribution of initial population, assume this disease-free equilibrium is achieved at $\boldsymbol{y}_\infty = (S^*, 0, 0, R^*)^\intercal$. Since

$$\boldsymbol{F} = \begin{bmatrix} 0 & \beta S^* \\ 0 & 0 \end{bmatrix}, \boldsymbol{V} = \begin{bmatrix} \mu_E & 0 \\ -\mu_E & \mu_I \end{bmatrix}, \boldsymbol{K} = \boldsymbol{F}\boldsymbol{V}^{-1} = \begin{bmatrix} \dfrac{\beta S^*}{\mu_I} & \dfrac{\beta S^*}{\mu_I} \\ 0 & 0 \end{bmatrix}$$

the basic reproduction number of this system at equilibrium is

$$\mathcal{R}_0 = \rho(\boldsymbol{K}) = \max\left\{0, \frac{\beta S^*}{\mu_I}\right\} = \frac{\beta S^*}{\mu_I}$$

The basic reproduction number for vector-borne disease can be defined similarly. Consider the vector-borne disease described in Figure 1.2. Assume the disease-free equilibrium is achieved at $\boldsymbol{y}_{H,\infty} = (S_H^*, 0, 0, R_H^*)^\intercal, \boldsymbol{y}_{V,\infty} = (S_V^*, 0, D_V^*)^\intercal$, and assume the model parameters for this system are

$$\boldsymbol{\beta}_{V\to H} = (\beta, 0, 0, 0)^\intercal, \boldsymbol{\beta}_{H\to V} = (\beta, 0, 0)^\intercal, \boldsymbol{\mu}_H = (0, \mu_E, \mu_I, 0)^\intercal, \boldsymbol{\mu}_V = (0, \mu_V, 0)^\intercal$$

Consider only the active stages in both human and vector populations, and arrange them

in the order of $\{E_H, I_H, I_V\}$. Then the next generation matrix is

$$\boldsymbol{F} = \begin{bmatrix} 0 & 0 & \beta S_H^* \\ 0 & 0 & 0 \\ 0 & \beta S_V^* & 0 \end{bmatrix}, \boldsymbol{V} = \begin{bmatrix} \mu_E & 0 & 0 \\ -\mu_E & \mu_I & 0 \\ 0 & 0 & \mu_V \end{bmatrix}, \boldsymbol{K} = \boldsymbol{F}\boldsymbol{V}^{-1} = \begin{bmatrix} 0 & 0 & \dfrac{\beta S_H^*}{\mu_V} \\ 0 & 0 & 0 \\ \dfrac{\beta S_V^*}{\mu_I} & \dfrac{\beta S_V^*}{\mu_I} & 0 \end{bmatrix}$$

Then the basic reproduction number of this system is

$$\mathcal{R}_0 = \rho(\boldsymbol{K}) = \max\left\{0, -\sqrt{\frac{S_H^* S_V^*}{\mu_I \mu_V}}\beta, \sqrt{\frac{S_H^* S_V^*}{\mu_I \mu_V}}\beta\right\} = \sqrt{\frac{S_H^* S_V^*}{\mu_I \mu_V}}\beta$$

1.5 Applications

In this section, we discuss two applications of our proposed general-purpose modeling framework in modeling the spread of Zika virus and avian influenza virus and demonstrate how the proposed models can assist in the effort of containing infectious disease outbreaks.

1.5.1 Modeling Zika Virus Outbreaks with Dynamic Human and Vector Populations

The Zika virus (ZIKV) outbreak in South American countries and its potential association with microcephaly in newborns and Guillain-Barré Syndrome led the World Health Organization to declare a Public Health Emergency of International Concern. To understand the ZIKV disease dynamics and evaluate the effectiveness of different containment strategies, we propose a compartmental model with a vector-host structure for ZIKV. The model utilizes logistic growth in human population and dynamic growth in vector population. Using this model, we derive the basic reproduction number to gain insight on containment strategies. We contrast the impact and influence of different parameters on the virus trend and outbreak spread. We also evaluate different containment strategies and their combination effects to achieve early containment by minimizing total infections. This result can help decision makers select and invest in the strategies most effective to combat the infection

spread. The decision-support tool demonstrates the importance of "digital disease surveillance" in response to waves of epidemics including ZIKV, Dengue, Ebola, and cholera. This section is adapted from the paper appeared in American Medical Informatics Association Proceedings 2016 [137].

Introduction

Zika virus (ZIKV) is an arbovirus belonging to the family *Flaviviridae* and is closely related to the dengue, yellow fever, Japanese encephalitis, and West Nile viruses. It was first isolated from a monkey in the Zika forest of Uganda in 1947; the second isolation was made from *Aedes africanus* mosquitoes in the same forest in 1948 [62]. Although earlier studies have suggested the possibility of human infection [193], before 2007, ZIKV rarely caused recognized "spillover" infections in humans, even in highly enzootic areas. In 2007, the first documented ZIKV outbreak was reported from Yap State, Federated States of Micronesia [73]. French Polynesia also reported an outbreak of 28,000 estimated cases in 2013 [40]. Since April 2015, Brazil has been experiencing a ZIKV outbreak. The outbreak has subsequently spread to other countries in South America, Central America, and the Caribbean [38, 39, 108].

ZIKV is primarily transmitted by the bites of infectious *Aedes* mosquitoes, especially *Aedes aegypti* and *Aedes albopictus* [100]. Non-vector borne transmissions, though rare, have also been reported [90]. Symptoms of ZIKV infection include fever, rash, conjunctivitis, and retro-orbital pain. These symptoms are usually mild and unspecific [112]. The ZIKV infection is difficult to detect since its symptoms are similar to those of dengue and chikungunya [100]; asymptotic cases are also frequent, which makes it even harder to precisely diagnose the ZIKV infection [105]. Although the symptoms of ZIKV infections are usually mild and unlethal, it has recently been connected to microcephaly in infants born to mothers infected with ZIKV during pregnancy, suggesting the possibility of intrauterine

infection [204]. ZIKV is also suspected to be linked to Guillain-Barré Syndrome (GBS) [172], a muscle weakness caused by the immune system damaging the peripheral nervous system.

With the ease of modern transportation, there is a potential that the ZIKV outbreak may become a global public health crisis [30]. Therefore, it is crucial to understand the transmission dynamics of ZIKV and establish a set of strategies that help to contain the outbreak effectively. Mathematical models have long been used to analyze infectious disease outbreaks to facilitate the development of policies and strategies for understanding disease propagation and containment strategies [101]. Since ZIKV is closely related to West Nile, dengue, and yellow fever viruses [130], it can be modeled with similar techniques. Because of dengue's global impact, its disease dynamics have been studied extensively, and compartmental models based on Susceptible, Infectious, Recovered (SIR), and Susceptible, Exposed, Infectious, Recovered (SEIR) hosts have been proposed and analyzed [61, 79–82, 84]. The population dynamics and life table of the major vector of ZIKV, Aedes mosquitoes, are also well studied [57, 78, 195, 203], and can be incorporated in the compartments for vectors.

In this study, we present a compartmental model for ZIKV based on SEIR host model and SEI vector model, using logistic growth in human population and dynamic vector population. We conduct sensitivity analysis on multiple parameters to understand the risk and transmission of ZIKV and evaluate the efficiency of different containment strategies. We discuss how this may assist in decision making to control and mitigate outbreaks.

Methods and Designs

Following the notation in the general-purpose model, the model for Zika virus consists of two parts: the compartments for human, subscripted with H; and the compartments for vec-

tors, subscripted with V. Although there have been a few reported deaths related to ZIKV, most ZIKV cases only result in mild symptoms among infectious individuals. We use an SEIR (Susceptible, Exposed, Infectious, Recovered) model to describe the disease spread among the human population, denoted as $S_H(t), E_H(t), I_H(t)$, and $R_H(t)$ respectively. The total human population at time t is given as $N_H(t) = S_H(t) + E_H(t) + I_H(t) + R_H(t)$. During the outbreak, the human population is assumed to follow a logistic growth with equal birth and death rates λ and carrying capacity K_H.

The compartments for vectors consist of both the infectious status and the life stages of the mosquitoes. A typical life cycle of a mosquito has four stages: egg, larva, pupa, and adult. Let $V_1(t), V_2(t)$ and $V_3(t)$ denote the number of eggs, larvae, and pupae of vectors in the environment. Since only female adults bite humans and transmit the virus, there are no infectious compartments for these three stages. Let $S_V(t), E_V(t)$ and $I_V(t)$ denote the number of susceptible, exposed, and infectious female adults in the environment at time t, and denote the total female vector population as $N_V(t) = S_V(t) + E_V(t) + I_V(t)$. Due to their short life cycle, the deceased stage of vector population is also considered; however, we do not include it in the analysis explicitly for simplicity. The transmission from susceptible to infectious in both human and vector population is associated with mosquito bites. Figure 1.3 shows the transition diagram.

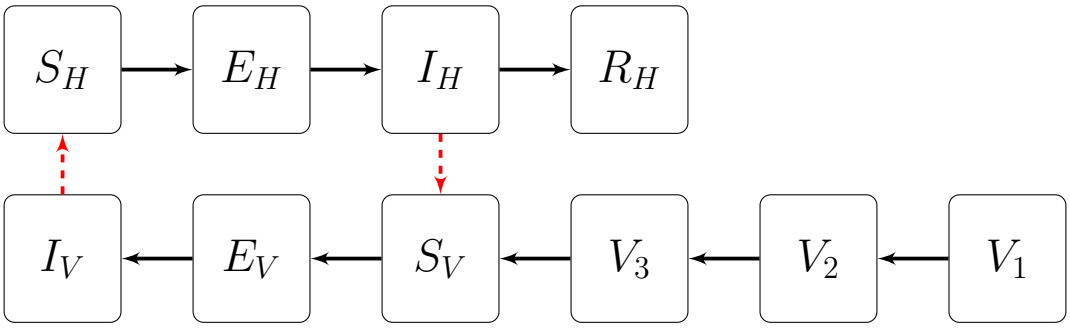


Figure 1.3: Zika: transition diagram of the compartmental model.

Let b denote the average biting rate per unit time of a female mosquito; b typically ranges from 0.3 to 0.5 per day [78]. Let β represent the percentage of bites that transmit ZIKV. In addition, let M denote the number of alternative hosts besides human in the environment. Therefore, following the same reasoning in deriving the general-purpose modeling framework for vector-borne disease, the exposure rate of a human to infection is $\beta b I_V/(N_H+M)$, and the exposure rate of a mosquito for infection is $\beta b I_H/(N_H+M)$. Assume the inverse incubation time of ZIKV for the human is α_H, and the recovery rate of an infectious individual is γ. Using the rule of logistic growth, the natural human population growth rate is λ, and the decreasing human population due to limited carrying capacity is $\lambda N_H^2/K_H$. This amount is distributed in the four compartments according to their relevant sizes. Plugging these parameters into Equation 1.10, the transitions between compartments for the human population can be described by the following system of ordinary differential equations:

$$\frac{\mathrm{d}}{\mathrm{d}t}S_H = \lambda N_H\left(1-\frac{S_H}{K_H}\right) - \frac{\beta b I_V}{N_H+M}S_H \qquad \frac{\mathrm{d}}{\mathrm{d}t}I_H = \alpha_H E_H - \gamma I_H - \lambda N_H\frac{I_H}{K_H}$$
$$\frac{\mathrm{d}}{\mathrm{d}t}E_H = \frac{\beta b I_V}{N_H+M}S_H - \alpha_H E_H - \lambda N_H\frac{E_H}{K_H} \qquad \frac{\mathrm{d}}{\mathrm{d}t}R_H = \gamma I_H - \lambda N_H\frac{R_H}{K_H}$$

For the vector compartments, let e_V denote the number of eggs laid per female mosquito per unit time, and α_V be the inverse incubation time of ZIKV in vectors. Let τ_i and μ_i be the reciprocal of the development time and mortality rate of vectors in stage $i \in \{1,2,3\}$ respectively. In addition, let μ_V denote the mortality rate of female adults. Assume q is the percentage of females among all adult vectors. Since mosquito larvae live in water with limited resources, the dynamics of the larvae population is also assumed to follow a logistic growth with environmental carrying capacity K_V. The mortality rate of larvae caused by limited carrying capacity is denoted as κ. Plugging these parameters into Equation 1.10, the transitions between compartments of vectors are described by the following system of

ordinary differential equations:

$$\frac{\mathrm{d}}{\mathrm{d}t}V_1 = e_V N_V - (\tau_1 + \mu_1)V_1 \qquad \frac{\mathrm{d}}{\mathrm{d}t}S_V = q\tau_3 V_3 - \frac{\beta b I_H}{N_H + M}S_V - \mu_V S_V$$
$$\frac{\mathrm{d}}{\mathrm{d}t}V_2 = \tau_1 V_1 - (\tau_2 + \mu_2)V_2 - \kappa\frac{V_2^2}{K_V} \qquad \frac{\mathrm{d}}{\mathrm{d}t}E_V = \frac{\beta b I_H}{N_H + M}S_V - (\mu_V + \alpha_V)E_V$$
$$\frac{\mathrm{d}}{\mathrm{d}t}V_2 = \tau_2 V_2 - (\tau_3 + \mu_3)V_3 \qquad \frac{\mathrm{d}}{\mathrm{d}t}I_V = \alpha_V E_V - \mu_V I_V$$

Refer to Appendix in this section to see the model matrices used for this case in Equation 1.10 to derive the systems of ordinary differential equations above. Table 1.2 summarizes the values of model parameters and their literature sources.

Table 1.2: Zika: parameter values in modeling the disease dynamics.

Name	Notation	Value	Literature Source
Number of infectious bites delivered per vector per day	βb	$0.05 - 0.12$	[78, 203]
Incubation time in human	$1/\alpha_H$	$2 - 4$ days	[24]
Incubation time in vector	$1/\alpha_V$	$8 - 12$ days	[108]
Recovery time in human	$1/\gamma$	$5 - 7$ days	[46]
Number of eggs laid per female mosquito per unit time	e_V	80/day	[57, 78]
Egg development time	$1/\tau_1$	3 days	[78, 195]
Egg mortality rate	μ_1	0.05	[78, 195]
Larvae development time	$1/\tau_2$	9 days	[78, 195]
Larva mortality rate	μ_2	0.025	[78, 195]
Pupa development time	$1/\tau_3$	3 days	[78]
Pupa mortality rate	μ_3	0.0025	[78]
Adult mortality rate	μ_V	0.05	[78, 195]
Female adult percentage	q	0.68	[195]

The system of ordinary differential equations above has a disease-free equilibrium if

$$e_V q\tau_1\tau_2\tau_3 > (\tau_1 + \mu_1)(\tau_2 + \mu_2)(\tau_3 + \mu_3)\mu_V$$

The number of noninfectious adult vectors at equilibrium if the inequality above holds is

$$S_V^* = \lim_{t\to\infty} S_V = \frac{e_v q\tau_1\tau_2\tau_3 - (\tau_1+\mu_1)(\tau_2+\mu_2)(\tau_3+\mu_3)\mu_V}{(\tau_1+\mu_1)(\tau_3+\mu_3)^2\mu_V^2\kappa} K_V q\tau_2\tau_3$$

The number of noninfectious humans at equilibrium follows the result of logistic growth, i.e.,

$$N_H^* = S_H^* = \lim_{t\to\infty} S_H = K_H$$

If we arrange the compartments corresponding to infectious stages as $\{E_H, I_H, E_V, I_V\}$, following the notation in deriving the formula for basic reproduction number, the components of the next generation matrix are given by

$$\boldsymbol{F} = \begin{bmatrix} 0 & 0 & 0 & \dfrac{\beta b S_H^*}{N_H^*+M} \\ 0 & 0 & 0 & 0 \\ 0 & \dfrac{\beta b S_V^*}{N_H^*+M} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \boldsymbol{V} = \begin{bmatrix} \alpha_H + \dfrac{\lambda N_H^*}{K_H} & 0 & 0 & 0 \\ -\alpha_H & \gamma + \dfrac{\lambda N_H^*}{K_H} & 0 & 0 \\ 0 & 0 & \mu_V + \alpha_V & 0 \\ 0 & 0 & -\alpha_V & \mu_V \end{bmatrix}$$

Therefore, the basic reproduction number for our model is the spectral radius of the next generation matrix $\boldsymbol{F}\boldsymbol{V}^{-1}$:

$$\mathcal{R}_0 = \rho(\boldsymbol{F}\boldsymbol{V}^{-1}) = \frac{\beta b}{K_H+M}\sqrt{\frac{\alpha_H\alpha_V K_H S_V^*}{(\alpha_H+\lambda)(\gamma+\lambda)(\alpha_V+\mu_V)\mu_V}}$$

To quickly contain the outbreak and reach a disease-free equilibrium, it is desirable to reduce the basic reproduction number such that $\mathcal{R}_0 < 1$. In that case, an infected individual produces less than one new infected individual during its infectious period and the outbreak will start to contain. Using the formula for $\mathcal{R}_0$, we may alter the model parameters by introducing interventions to contain the outbreak. In particular, the following two strategies are considered:

1. Reduce biting rate b by avoiding mosquito bites, using insect repellents, wearing

long-sleeved clothes and long pants, and using air conditioning and window/door screens to keep mosquitoes outside.

2. Reduce adult female vector population or increase adult female vector mortality rate μ_V at equilibrium by widely applying insecticides in areas with high (infectious) mosquito population, or introducing genetically modified *Aedes* mosquitoes, which will cause its offspring to die by reducing larval survival rate and adult longevity [21, 44].

To validate the performance of the model and investigate the effectiveness of containment strategies, we perform sensitivity and scenario analysis using the 2015 ZIKV data. For sensitivity analysis, we vary the recovery time (γ) and the incubation time of human (α_H). For containment strategies, we experiment with biting rate (b), adult vector mortality rate (μ_V), carrying capacity of larvae (K_V), and adult vector population (N_V). We report the estimated number of infections and time to reach containment. The simulation is performed until an asymptotic behavior is achieved. This signals containment is reached – the cumulative number of infections plateaus.

Results

Table 1.3 summarizes the parameters used in modeling the Zika disease dynamics for Rio de Janeiro State (Sentinel surveillance sites reports [33]) and for a hypothetic town with 100,000 people. Between January 2015 to July 2015, a total of 364 suspected cases were reported in Rio de Janeiro State. By November, the outbreak was contained with approximately 440 infection cases.

Rio de Janeiro State, Brazil Figure 1.4 contrasts the reported cases versus simulation by varying the recovery time (γ) and the incubation time of human (α_H). The results suggest that Table 1.3 provides good estimates for the reported cases (green solid curve vs red dotted curve). When the recovery time increases from 6 to 8 days (dotted vs dashed

Table 1.3: Zika: estimated parameter values used in the simulations.

Name	Notation	Rio de Janeiro State	A town of 100,000
Initial human population	$N_H(0)$	16,231,365	100,000
Initial human infections	$I_H(0)$	1	10
Initial adult vector population	$N_V(0)$	70,000,000	500,000
Initial exposed adult vectors	$E_V(0)$	5	1,000
Initial infectious adult vectors	$I_V(0)$	5	1,000
Initial number of eggs	$V_1(0)$	180,000,000	500,000
Initial number of larvae	$V_2(0)$	90,000,000	150,000
Initial number of pupae	$V_3(0)$	70,000,000	100,000
Infectious daily biting rate	βb	0.071	0.09
Carrying capacity for human	K_H	30,000,000	120,000
Carrying capacity for larvae	K_V	100,000,000	5,000,000
Larvae mortality rate due to limited carrying capacity	κ	1	1
Number of alternative hosts	M	100,000	80,000
Human population growth rate	λ	0.00005	0.00005
Recovery time of human	$1/\gamma$	7 days	7 days
Incubation time of human	$1/\alpha_H$	3 days	3 days
Incubation time of vector	$1/\alpha_V$	10 days	10 days
Basic reproduction number (estimate from parameters)	$\mathcal{R}_0$	1.1139	1.109
Simulation start date		Jan 1, 2015	Day 0
Implementation of strategies		May 1, 2015	Day 100
Time for containment without intervention		Nov, 2015	Day 900

green curves), the cumulative infections increases from 340 to 566 by July 2015 (66.5% increase). Longer recovery time also leads to delayed containment. This confirms that recovery time has a significant impact on disease containment. Thus timely treatment and vaccination for combating ZIKV are critical. Varying the human incubation period from 2 days to 4 days has a relatively minor influence on the overall infections (from 427 to 459, 7.5% increase) and the outbreak trend (contrast the three solid lines).

Figure 1.5 contrasts containment effectiveness for four strategies: reducing biting rate (b), increasing adult vector mortality rate (μ_V), reducing larvae carrying capacity (K_V),

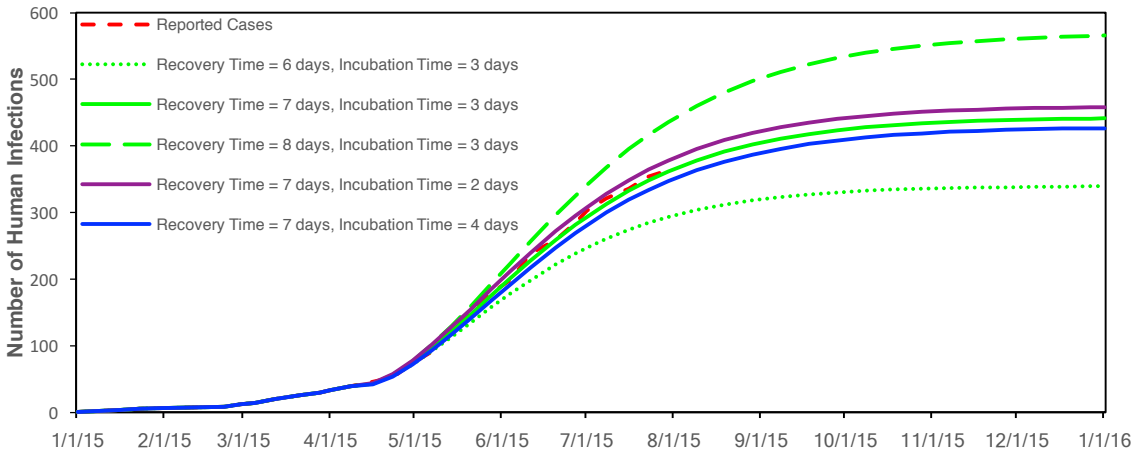
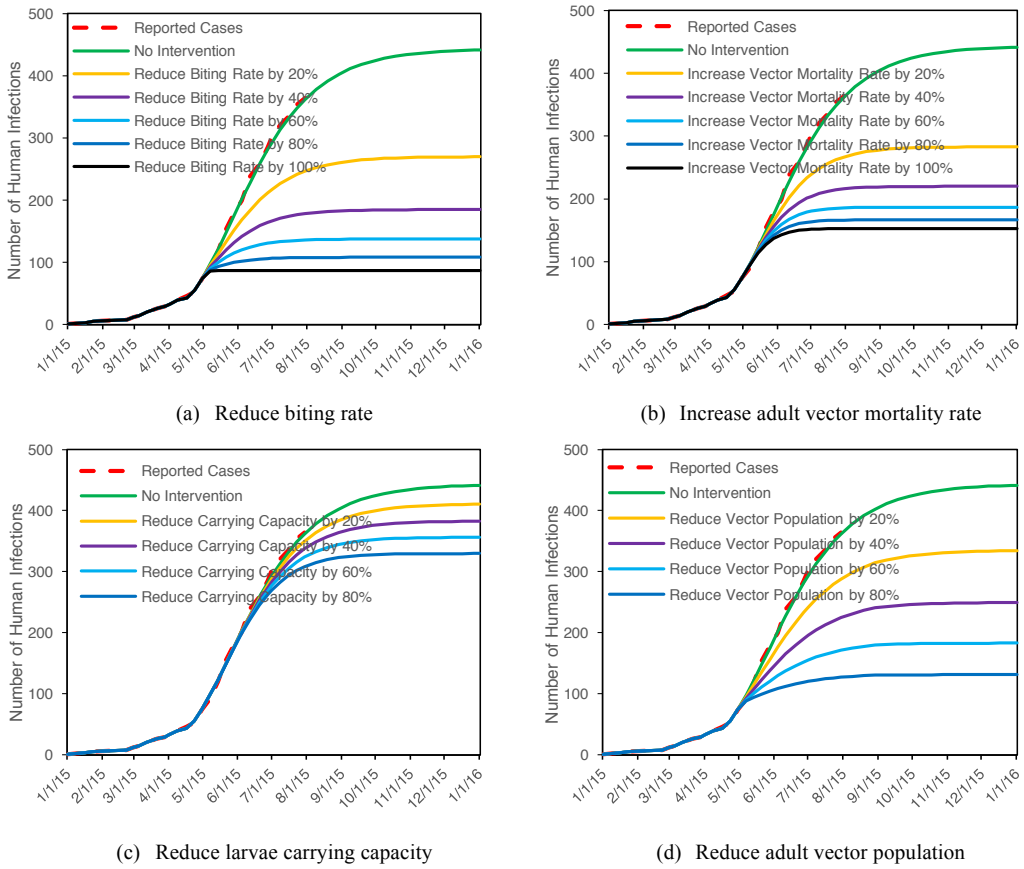


Figure 1.4: Zika: sensitivity analysis for Rio de Janeiro State.

and reducing adult female vector population (N_V). Without these strategies, total infection reaches 443 and will be contained by October 2015. In our analysis, we assume each strategy is implemented on May 1, 2015. Figure 1.5(a) shows that reducing biting rate by a mere 20% will lead to an early containment by August 2015 (90 days after implementation) with 270 total infections (39% reduction). The containment is achieved rapidly (¡ 60 days after implementation) with total infection under 200 when biting rate is reduced by at least 40%. This proves that reducing biting rate is highly effective. This strategy is also relatively easy to implement. Increasing the mortality rate of adult vectors is also effective. Figure 1.5(b) shows that increasing the vector mortality rate by 20% will lead to an August containment with total infection of 282 (36% reduction). Containment will be achieved instantly with no more than 220 infections if the mortality rate is increased by at least 40%. While this strategy can be achieved by introducing genetically modified mosquitoes to the environment, it is difficult to realize and economically inferior when compared to other strategies.

The larvae carrying capacity has to be reduced by 80% in order for total infection to reduce to 330 (26% reduction) with containment achieved by October (Figure 1.5(c)).



(a) Reduce biting rate

(b) Increase adult vector mortality rate

(c) Reduce larvae carrying capacity

(d) Reduce adult vector population

Figure 1.5: Zika: contrast of outbreak trends for four strategies in Rio de Janeiro State.

Nonetheless, examining and clearing the water ponds and humid areas where the larvae live remains an important strategy for public health protection. Reducing vector population not only affects the number of infections at containment, it also changes the trend of the outbreak (Figure 1.5(d)). Specifically, reducing total adult vector population by 20% would postpone the outbreak by a month and reduce total infection to 334. Containment can be rapid when the reduction is targeted at 80% (total infection reduces to 130). Figures 1.5(c) and 1.5(d) highlight that reducing adult population is more effective, though applying insecticide widely may come with numerous environmental and health issues.

In practice, combination strategies are often most effective. For brevity, Figure 1.6 reports containment results for the combination of two strategies. In both figures, the vertical

axis represents the cumulative number of human infections at containment. When 40% reduction is achieved in both biting rate and larvae carrying capacity; it results in 180 total infections (Figure Figure 1.6(a)). Similarly, the combination of reducing vector population by 20% and increasing vector mortality rate by 20% results in 232 total infections (Figure 1.6(b)).

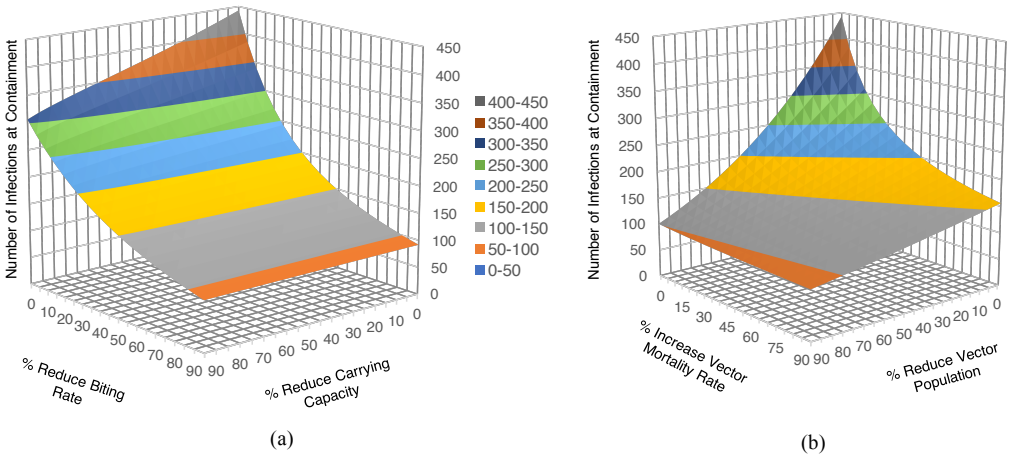
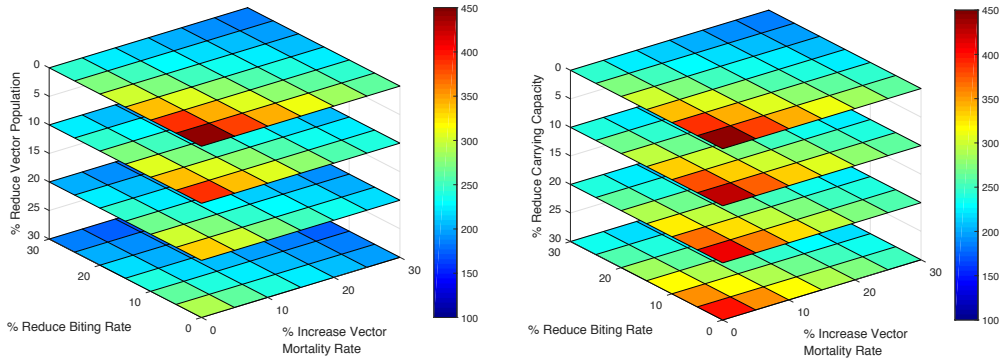


Figure 1.6: Zika: infections with combinations of 2 strategies in Rio de Janeiro State.

Figure 1.7 shows the containment results for the combination of three strategies, where the color of each grid represents the cumulative human infections at equilibrium. For example, if increasing vector mortality rate, reducing biting rate, and reducing adult vector population are all carried out at 10% level, the total infections will be 250 (Figure 1.7(a)). Since it can be difficult to achieve high effectiveness using a single strategy, combination strategies are appealing. Our results also display marginal improvement beyond certain threshold values in the strategies. This helps policymakers to invest their resources optimally for best possible gain.

A hypothetic town with 100,000 people To better understand the behavior and sensitivity of our model and the disease spread characteristics, we perform the analysis on a town with 100,000 people, using the parameters given in Table 1.3. Without any inter-



(a) (b)

Figure 1.7: Zika: infections with combinations of 3 strategies in Rio de Janeiro State.

vention, containment will occur at Day 900 with 42,000 overall infections. We observe some similarities as in the Rio de Janeiro State analysis. An increase in human recovery time delays the containment and increases the total number of infections (Figure 1.8(a)). Similarly, the incubation period for humans has little effect on the overall disease trend (Figure 1.8(c)). On the other hand, the incubation time for the vectors significantly affects the overall number of infections while exhibiting the same outbreak pattern (Figure 1.8(d)). We explore alternative hosts and their impact on the pandemic. Increasing the number of alternative hosts from 20,000 to 100,000 (Figure 1.8(b)) will slow the outbreak and reduce the total number of infections by 83.3%. This suggests that introducing alternative hosts has potential in mitigating an outbreak. However, it is possible that these hosts can get infected with ZIKV and spread it to humans via other means. The ecological and disease impact cannot be fully assessed at this point, thus rendering it an unsafe venue for ZIKV disease containment.

We assume each strategy is implemented on Day 100 after the outbreak. Figure 1.9 reports containment effectiveness for each strategy. Reducing bite rate by 20% (Figure 1.9(a)) and increasing vector mortality by 20% (Figure 1.9(b)) have similar effectiveness in containing the disease spread (Day 500 vs Day 300) and with comparable overall infection (12,000 vs 11,000). Both strategies are very effective - reducing infections by over 70%.

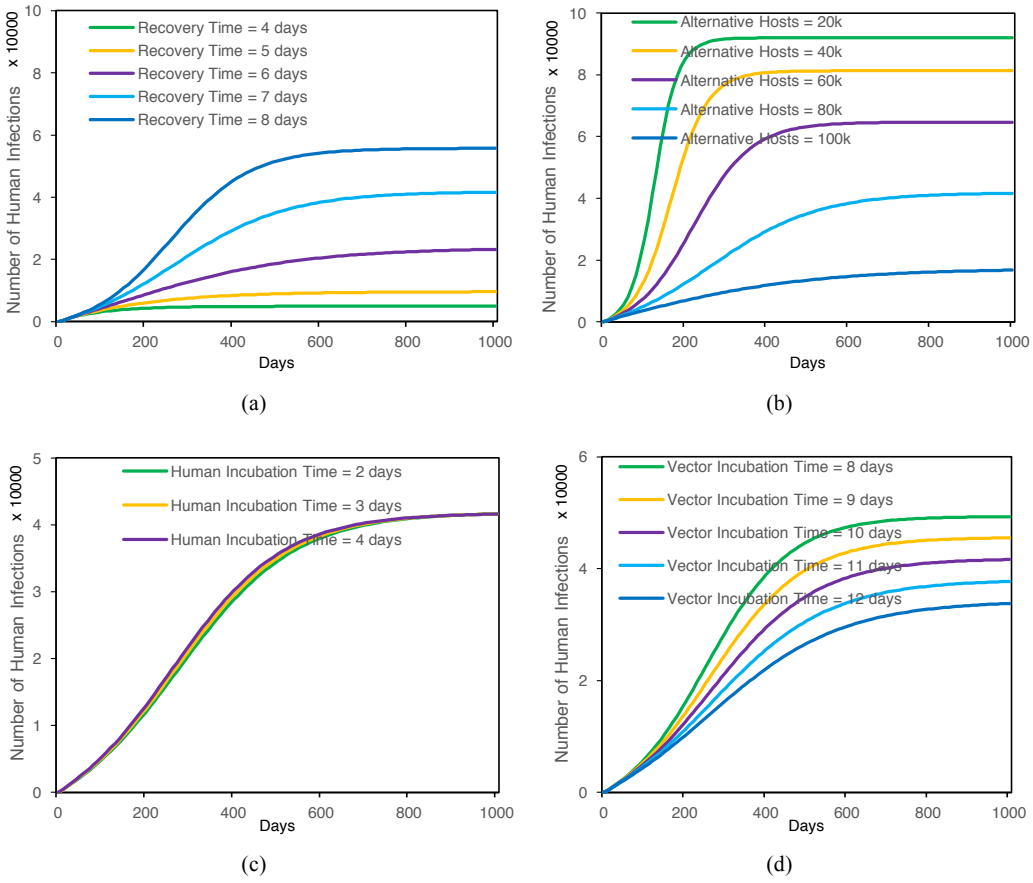
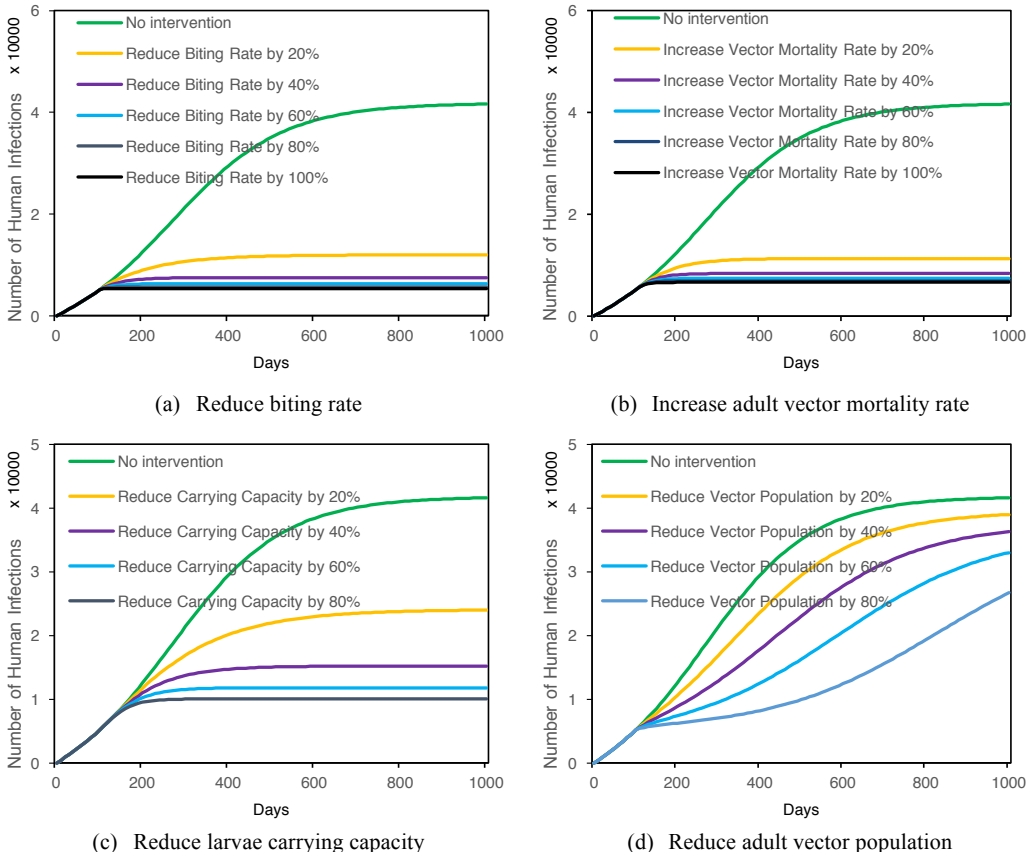


Figure 1.8: Zika: sensitivity analysis for the hypothetic town.

Reducing larvae carrying capacity has a significant impact on the outbreak (Figure 1.9(c)). In particular, over 40% reduction in infection is observed when 20% reduction in carrying capacity is in effect. On the other hand, reducing vector population is not as effective in containing the ZIKV outbreak (Figure 1.9(d)). Nonetheless, reducing vector population does influence the pandemic course: when the vector population is reduced by 20%, it postpones the onset of the outbreak by approximately 100 days and reduces the total infection by 9.5%. These results are different from the findings in the Rio de Janeiro State analysis where reducing the carrying capacity of larvae proves to be not very effective. Similarly, combination strategies can be used in this scenario to achieve better containment results.



(a) Reduce biting rate

(b) Increase adult vector mortality rate

(c) Reduce larvae carrying capacity

(d) Reduce adult vector population

Figure 1.9: Zika: contrast outbreak trends for four strategies in the hypothetic town.

Discussion

This work demonstrates the key importance of "digital disease surveillance" in response to waves of epidemics including ZIKV, Dengue, Ebola, and cholera. Specifically, we develop a compartmental model with a vector-host structure utilizing logistic growth for human population and a dynamic vector population describing its full life cycle. We analyze the development of ZIKV outbreak and explore key parameters that influence epidemic severity. We study the effect of average recovery time (of human) from infection, the number of alternative hosts, and the ZIKV incubation time in both vectors and humans. The analyses are performed on two populations with real data from Rio de Janeiro State. Our results reveal that the outbreak is sensitive to the first three parameters while human incubation period has only a marginal impact.

We examine four containment strategies: reducing biting rate, increasing vector mortality rate, reducing larvae carrying capacity, and reducing vector population. All four strategies are effective (in varying degrees) in containing the outbreak and reducing the overall infection. While findings for reducing biting rate and increasing vector mortality rate are similar for both populations (both are superior to other strategies), results for the larvae carrying capacity and reduction in vector population differ. These findings suggest environmental and demographic information should be considered when determining proper containment strategies. Our study also shows that reducing vector population early on can help delay the onset of the outbreak.

Combination strategies are both promising in practice and cost-effective in achieving early containment. The model provides a decision support framework for policymakers to estimate the cost-effectiveness for each prevention measure. Public health departments should select a strategy portfolio compatible with their local environment and regional demographics. In addition, the public should be educated and informed of ZIKV status. Population behaviors (protecting themselves by reducing biting, cleaning water ponds to rid of larvae, etc.) have a demonstrably significant impact on containing and mitigating the outbreak.

The multiple components involving the dynamics of human and vector populations in this model allow flexibility in characterizing disease spread and performing strategic analysis. However, obtaining/determining all essential input parameters for this to be practical may prove difficult. The model can be simplified without diminishing its quality and rigor in disease dynamics and infection/containment outcome prediction. Such a model may be desirable when detailed input data are not readily available. For example, the natural human population can be assumed to be constant throughout the outbreak since the magnitude of natural population growth is relatively small within the pandemic period. The mosquito

population life-cycle can also be condensed to a simple birth-death process. we propose a simplified model with constant human and mosquito population to study the ZIKV outbreak. The study predicts the disease trends and characterizes containment strategies for Brazil and Puerto Rico. The model has accurately predicted ZIKV outbreak trends for these two countries. It predicts the number of infectious pregnant women in Puerto Rico within 5% of the actual documented numbers (as of July 2016). Although the simplified model is statistically useful and usable, additional components and parameters presented herein should help decision-makers better understand the details of ZIKV outbreaks. It may also provide them with more analytic-based options in deciding the proper containment strategies when such input data are available.

Modern information technologies which are capable of tracking disease spread and vector activities play an indispensable role in supporting disease modeling and in mitigating and containing epidemiology outbreaks. Informatics applications such as social media and internet-based participatory surveillance systems have already demonstrated their capability in helping us understand and prevent disease outbreaks like Dengue and influenza [154, 211]. Software and hardware such as ArcGIS platform and mobile DNA sequencing lab are also designed to help track the current ZIKV outbreak. These technologies capture valuable inputs for our model, which is very flexible and can be easily adapted in real-time as real data feeds arrive. With the support of these technologies, the model parameters can be estimated accurately and updated on a regular basis. The model output can then be used to predict the status of the outbreak in a timely manner. It also reflects any changes in the environment, policy, and human and social behavior. Conversely, the model can also serve as a guideline for the application of these decision technologies such that they are deployed and utilized optimally by policymakers. Informatics can help us understand the underlying mechanism of the pandemic spread. It also enables decision-makers and policymakers to understand/update the planning and response strategy in real time to maximize the contain-

ment effect while minimizing the overall infection/mortality and the associated costs and health burden.

There are limitations to our study. We assume a homogeneous human population in the model, which may not hold in practice. The possibility of contacting mosquitoes and the chance of getting infected from an infectious vector may differ among individuals and the environment. This can be overcome by creating additional compartments for each homogeneous subgroup. We have incorporated this subgroup concept in our vaccine prioritization analysis with great success [143]. However, this requires even more parameter estimation and may lead to over-fitting. In addition, our conclusions are strongly influenced by results from a particular demographic and geographic profile. Local decision makers should input proper (best estimate) parameters into the model in order to conduct accurate analyses.

The current model does not include treatment or vaccination effects since they are currently under developed. With a deeper understanding of ZIKV and the presence of medications and vaccines, these factors can be incorporated under our modeling framework.

Appendix

Following the notations in our general-purpose modeling framework for vector-borne diseases, if we align the order of human compartments as S_H, E_H, I_H, R_H and align the order of vector compartments as $V_1, V_2, V_3, S_V, E_V, I_V, D_V$, the model matrices for the human and vector populations are

$$\boldsymbol{D}_H = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \boldsymbol{C}_{V\to H} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \boldsymbol{\mu}_H = \begin{bmatrix} 0 \\ \alpha_H \\ \gamma \\ 0 \end{bmatrix}$$

$$\boldsymbol{\beta}_{V\to H} = \begin{bmatrix} \beta b/(N_H+M) \\ 0 \\ 0 \\ 0 \end{bmatrix}, \lambda(\boldsymbol{y}_H) = \frac{\lambda N_H}{K_H} \begin{bmatrix} K_H - S_H \\ -E_H \\ -I_H \\ -R_H \end{bmatrix}$$

$$\boldsymbol{D}_V = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \boldsymbol{L}_V = \begin{bmatrix} 0 & 0 & 0 & e_V & e_V & e_V & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\boldsymbol{R}_V = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{\tau_1}{\tau_1+\mu_1} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{\tau_2}{\tau_2+\mu_2+\kappa V_2/K_V} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{q\tau_3}{\tau_3+\mu_3} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{\mu_1}{\tau_1+\mu_1} & \frac{\mu_2+\kappa V_2/K_V}{\tau_2+\mu_2+\kappa V_2/K_V} & \frac{(1-q)\tau_3+\mu_3}{\tau_3+\mu_3} & 1 & 1 & 1 & 0 \end{bmatrix}$$

$$\boldsymbol{C}_{H\to V} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \boldsymbol{\mu}_V = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \alpha_V \\ 0 \\ 0 \end{bmatrix}, \boldsymbol{\beta}_{H\to V} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \beta b/(N_H + M) \\ 0 \\ 0 \\ 0 \end{bmatrix}, \boldsymbol{\tau}_V = \begin{bmatrix} \tau_1 + \mu_1 \\ \tau_2 + \mu_2 \\ \tau_3 + \mu_3 \\ \mu_V \\ \mu_V \\ \mu_V \\ 0 \end{bmatrix}$$

Plugging the model matrices above into the modeling framework presented in Equation 1.10 results in the system of ordinary differential equations used in the model for Zika virus outbreak.

1.5.2 Modeling Avian Influenza Outbreaks with Migration and Transportation

Avian influenza viruses have caused infections and deaths in wild birds, commercial poultries, and humans. It poses an increasing threat of a pandemic. To understand the transmission dynamics of avian influenza viruses and assess the effectiveness of different containment strategies, we develop a flexible modeling framework based on multi-layer compartmental models for digital disease surveillance and response in combating pandemics. The model can accommodate other disease outbreaks under diverse settings. We demonstrate its usage on avian influenza and derive the basic reproduction number and spread characteristics. We contrast the effectiveness of different containment strategies and their combination effect in protecting both the human and the bird population. This section is adapted from the paper appeared in American Medical Informatics Association Proceedings 2017.

Introduction

Avian influenza is caused by infection with avian influenza Type A viruses. These viruses were first isolated from common terns in South Africa [25] and was later extensively stud-

ied. They occur naturally among wild birds and can infect domestic poultry and other animal species, including cats and swine. The highly pathogenic avian influenza (HPAI), the very virulent types of avian influenza Type A viruses (majorly subtypes H5 and H7), can cause almost 100% death for birds [5]. They do not normally infect humans, but sporadic human infections with avian flu viruses have been reported, with approximate 50% mortality rate [126]. There is no clinical evidence that avian influenza can be directly transmitted between humans, but such possibility remains through mutant avian influenza viruses and should be closely monitored to reduce the infections and mortalities among humans [191, 214].

Primary outbreaks of HPAI in poultry have been reported multiple times. The more recent outbreaks started in 2003 when multiple avian influenza cases in wild birds, domestic poultry, and humans were reported worldwide. The H5N1 influenza outbreak in Southeast Asia in 2003 infected more than 130 people in Vietnam, Thailand, and Cambodia and killed more than half of them [207]. The same subtype appeared in wild birds and domestic poultry in China afterwards, causing human infections as well [128]. It was estimated that by 2005, this outbreak had caused more than $10 billion in losses to the Southeast Asian poultry industry [23]. Large outbreaks in commercial poultry farms in industrialized countries in Europe and the United States have also been reported [93], indicating that avian influenza is not only a threat to human health, it also jeopardizes the poultry industry and has a significant impact on the economy. Predicting the trend of an avian influenza outbreak and implementing an optimal intervention strategy to achieve early containment is crucial. Pandemic planning and large-scale surveillance are essential in promoting effective responses [58]; with real-time analysis of surveillance data being critically important to detect and predict an outbreak in a timely manner.

Numerous mathematical models have been developed to describe the spread of an epi-

demic to facilitate decision making using surveillance data. During the past two decades, various models were proposed for general influenza and avian influenza. A probability-based quantitative analysis was used to estimate the risk of an emerging pandemic strain of avian influenza [86]. Agent-based simulation of individual and community activities was developed to prevent and control influenza pandemics [43]. Stochastic models were developed to predict the worldwide spread of pandemic influenza based on global travel information [52]. Compartmental models remain a popular method for pandemic studies. Ordinary differential equation-based methods were used to interpret the mutation process of avian influenza [115], model the dynamics of human infections [61], and investigate nonlocal epidemics [3].

In this section, we use the general-purpose disease modeling framework for vector-borne diseases to model the spread of avian influenza virus. We use the avian influenza outbreak data from Egypt and United State to tune the model parameters, and perform sensitivity analysis to understand the risk of infection, and evaluate the efficiency of different containment strategies in terms of both human and bird infections.

Methods and Designs

The model for the spread of avian influenza is derived from the general form of the vector-borne disease, where the compartments for human are subscripted by H, and the compartments for birds (vectors) are subscripted by B. The infections among birds are associated with contacting infectious birds, while the transmission from the susceptible to the infectious in human is associated with contacting infectious birds, as well as their saliva, mucous, and feces. The spread of avian influenza A viruses from one ill human to another is rare. However, because of the high possibility of viruses to mutate and gain such an ability, the human-to-human transmission is also included in the model.

In this study, we apply an SEIRD (Susceptible, Exposed, Infectious, Recovered, Deceased) model to describe the disease spread among the human population. The human population is assumed to follow a logistic growth with carrying capacity K_H and growth rate r during the outbreak. The virus transmission among birds is modeled with an SID (Susceptible, Infectious, Deceased) model. The bird population is assumed to have a constant death rate ν and a variable birth rate $\Lambda(t)$ such that the total population is constant K_B. Figure 1.10 shows the transition diagram.

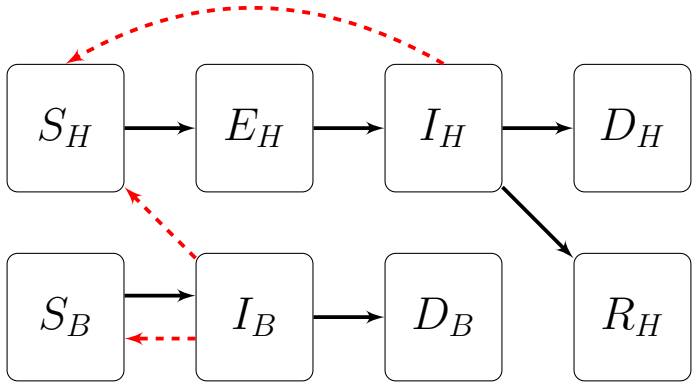


Figure 1.10: Avian influenza: transition diagram of the compartmental model.

Assume the probability of a human recovering from an infection is p, the baseline contact rate between humans is β_H, and the baseline contact rate between humans and birds is $\beta_{B\to H}$. Under this assumption, if we align the human compartments as S_H, E_H, I_H, R_H, D_H, and align the order of bird compartments as S_B, I_B, D_B, then the model matrices for the human population are

$$\boldsymbol{D}_H = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & p & 0 & 0 \\ 0 & 0 & 1-p & 0 & 0 \end{bmatrix}, \boldsymbol{C}_{H\to H} = \begin{bmatrix} 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \boldsymbol{C}_{B\to H} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\boldsymbol{\beta}_{B\to H}=\begin{bmatrix}\beta_{B\to H}\\0\\0\\0\\0\end{bmatrix},\boldsymbol{\beta}_{H\to H}=\begin{bmatrix}\beta_H\\0\\0\\0\\0\end{bmatrix},\boldsymbol{\mu}_H=\begin{bmatrix}0\\\mu_{HE}\\\mu_{HI}\\0\\0\end{bmatrix},\lambda(\boldsymbol{y}_H)=\frac{rN_H}{K_H}\begin{bmatrix}K_H-S_H\\-E_H\\-I_H\\-R_H\\0\end{bmatrix}$$

Assume the baseline contact rate between birds is β_B. Then the model matrices for the bird compartments are

$$\boldsymbol{D}_B=\begin{bmatrix}0&0&0\\1&0&0\\0&1&0\end{bmatrix},\boldsymbol{C}_{B\to B}=\begin{bmatrix}0&-1&0\\0&1&0\\0&0&0\end{bmatrix},\boldsymbol{\mu}_B=\begin{bmatrix}0\\\mu_{BI}\\0\end{bmatrix}$$

$$\boldsymbol{\beta}_{B\to B}=\begin{bmatrix}\beta_B\\0\\0\end{bmatrix},\lambda(\boldsymbol{y}_B)=\begin{bmatrix}\max\{K_B-S_B-I_B,0\}-\nu S_B\\-\nu I_B\\0\end{bmatrix}$$

By arranging the compartments corresponding to infections as E_H, I_H, I_B and evaluating the components of next generation matrix at equilibrium, we have

$$\boldsymbol{F}=\begin{bmatrix}0&\beta_{H\to H}K_H&\beta_{B\to H}K_H\\0&0&0\\0&0&\beta_{B\to B}K_B\end{bmatrix},\boldsymbol{V}=\begin{bmatrix}\mu_{HE}+r&0&0\\-\mu_{HE}&\mu_{HI}+r&0\\0&0&\mu_{BI}+\nu\end{bmatrix}$$

By applying the method of next generation matrix, the basic reproduction number for the system is

$$\mathcal{R}_0=\rho(\boldsymbol{F}\boldsymbol{V}^{-1})=\max\left\{\frac{\beta_{H\to H}K_H\mu_{HE}}{(\mu_{HE}+r)(\mu_{HI}+r)},\frac{\beta_{B\to B}K_B}{\mu_{BI}+\nu}\right\}$$

The first component in the equation is the basic reproduction number within the human compartments, and the second component is the basic reproduction number within the bird

compartments. Since human-to-human infection of avian influenza is inefficient and rare [26], we can conclude that $\beta_{H\to H}K_H \ll \beta_{B\to B}K_B$, thus the basic reproduction number of the system is determined by the infections among the birds. To quickly contain an avian influenza outbreak, it is desirable to reduce the basic reproduction number such that $\mathcal{R}_0 < 1$ so that an infected bird produces less than one new infection during its infectious period. This can be achieved by either reducing the adjusted baseline contact rate $\beta_{B\to B}$ by quarantining the poultry farms with confirmed cases, or reducing the bird population K_B by slaughtering infectious birds, both wild and live poultry.

To examine the performance of the model and understand the effectiveness of containment strategies, we perform sensitivity and scenario analysis using the 2010 avian influenza data in Egypt. We vary the contact rate between birds ($\beta_{B\to B}$), and the duration of infection in birds (μ_{BI}). We report the estimated number of infections in both birds and humans, as well as time to reach containment. The simulation is performed until an asymptotic behavior is achieved, i.e., when no new cases are reported.

Results

Egypt. Egypt is considered an important poultry producer in the Arab world with 1.5 million permanent workers and 1 million temporary workers in poultry production. The combination of commercial poultry, backyard birds, and live-bird markets exposes Egypt to the danger of avian influenza outbreak. Evidence shows that stable lineages of H5N1 viruses have been established among chickens and humans in Egypt, and the control efforts and international cooperation is in urgent need [2]. We use the data of reported infections in live poultry of an HPAI outbreak in Egypt during January to June 2010 for our analysis. The total number of reported cases in live poultry was 317,400 by containment.

Figure 1.11 contrasts the reported cases in live poultry against the simulation result

Table 1.4: Avian influenza: parameters used in modeling the outbreak in Egypt.

Name	Notation	Value	Literature Source
Initial live poultry population	S_B	1.8×10^9	[2]
Initial infected live poultries	I_B	3,000	Fitted
Average live poultries lifespan	$1/\nu$	60 days	[74]
Average live poultries lifespan after infection	$1/\mu_{BI}$	10 days	[114]
Baseline infection rate between live poultries	$\beta_{B\to B}K_B$	58/day	Fitted
Baseline infection rate between live poultries and humans	$\beta_{B\to H}K_H$	2×10^{-4}/day	[114]
Baseline infection rate between humans	$\beta_{H\to H}K_H$	1.5×10^{-7}/day	[114]
Initial human population (Egypt)	S_H	8.2×10^7	World Bank
Natural population growth rate (Egypt)	r	1.6%	World Bank
Basic reproduction number	$\mathcal{R}_0$	1.362	Calculated

by varying the contact rate ($\beta_{B\to B}K_B$) and the duration of infection (μ_{BI}) in live poultry. The results suggest that Table 1 provides good estimates for the reported cases (green solid curve from simulation against the red dotted curve of reported cases). When the baseline contact rate increases from 56 per day to 58 per day (dotted against the dashed green curves), the cumulative infections in bird compartments increase from 0.22 million to 0.46 million by June 2010 (110% increase). Varying the duration of infection in birds seems to have a larger impact on the trend of the outbreak (the three solid lines). When increasing the bird infection time from 9 days to 11 days, not only the total infections at containment sees a 4-fold increase (0.16 million to 0.6 million), the time to containment is delayed by two months as well (April 2010 to June 2010). Therefore, reducing the infection duration of birds plays a crucial role in controlling the trend of the avian influenza outbreak. This can be achieved by providing medical countermeasures to affected poultry farms, or quickly removing/isolating infectious live poultry upon identification of an outbreak.

Next, we focus our analysis on the infections among humans using the parameters from Table 1.4 and investigate the effectiveness of three containment strategies: 1) reducing

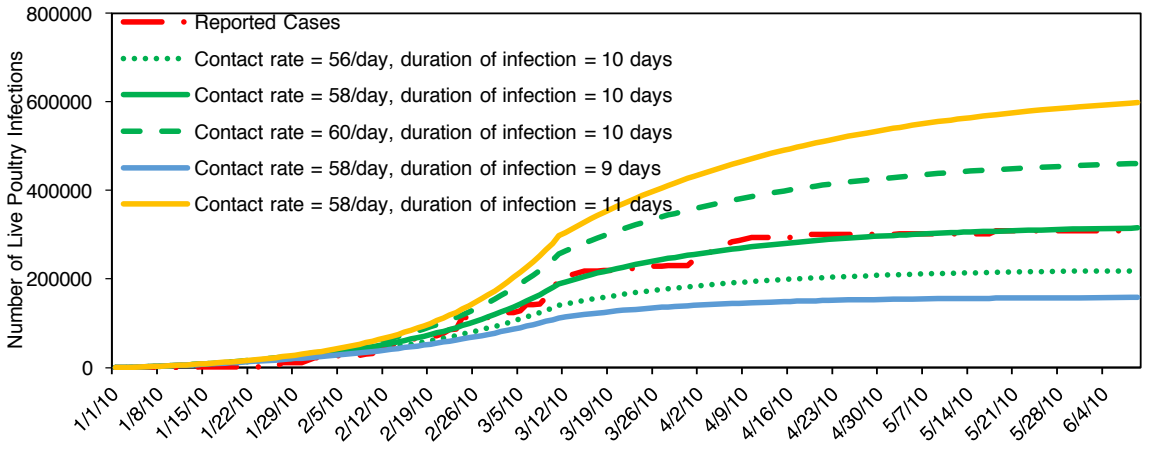


Figure 1.11: Avian influenza: sensitivity analysis of the outbreak in Egypt.

contact rate to live poultry ($\beta_{B\to H}$), 2) slaughtering infectious live poultry (I_B), and 3) quarantining infectious humans (I_H). Each strategy is assumed to be in effect starting on February 20, 2010. Without any intervention, the total infection reaches 540 and will be contained by June 2010.

Figure 1.12(a) shows that reducing contact rate by 20% will reduce the total infections to 453 (16% reduction), although it does not significantly improve the time of containment. However, the containment will be achieved by May 2010 with 278 infections when the contact rate is reduced by 60%. This shows that reducing contact rate is highly effective in minimizing the total infections and containment time. This strategy is also easy to implement: residents should be advised to avoid contacting birds or their saliva, mucous and feces; and poultry farm workers should be advised to use proper protective equipment while working close to the poultry.

Figure 1.12(b) suggests that slaughtering infectious live poultry is equally effective. It has similar efficient frontiers as reducing contact rate in controlling human infections. However, it also controls the outbreak within the live poultry. Unlike reducing contact rate, which works totally independent of the bird compartments, reducing infectious poultry

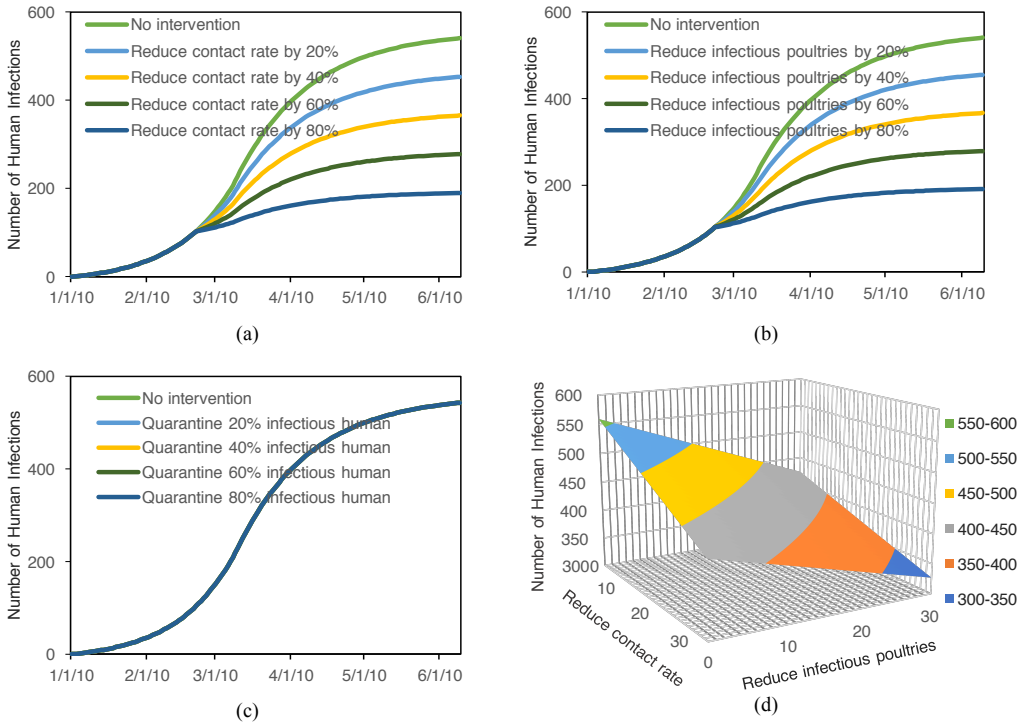


Figure 1.12: Avian influenza: contrast outbreak trends for different strategies in Egypt.

also reduces the danger of susceptible birds being exposed to the virus. This leads to earlier containment within the live poultry. However, there are practical difficulties with this approach: the management of the bird carcasses and the potential economic impact. Incineration of the dead birds has been a common practice.

Figure 1.12(c) shows that with the current baseline infection among humans at 1.5×10^{-7}. The effect of human-to-human infection is negligible: all 5 curves are almost the same. But this could become a threat due to the high mutation probability of avian influenza viruses. Mutant species which can be easily transmitted among humans may emerge; hence epidemiologists and healthcare leaders should closely monitor such a possibility.

Figure 1.12(d) shows the combination effect of reducing contact rate and reducing infectious poultry. When both strategies are implemented at 20% level, the total number of infections at containment is approximately 400. Since implementing a single strategy at a

high efficiency level may be inherently difficult, combination strategies are appealing and can generate satisfactory containment results.

A significant concern to public health officials is to understand when the intra-human infection of avian influenza will pose a major threat to public health. By varying the baseline infection rate between humans ($\beta_{H\to H}K_H$) while keeping other parameters fixed as given in Table 1.4, Figure 1.13 shows that the trend of the avian influenza outbreak in human remains insignificant unless the baseline infection rate is increased from 1.5×10^{-7} to 0.15. Although such a possibility is small, it will result in more than 5,400 infections in humans at containment, 7 times more when compared to the original scenario. Since the human mortality rate is very high, it is imperative that epidemiologists and healthcare officials are vigilant in monitoring the mutations of avian influenza viruses, and take proper and timely actions should such high infectivity mutations occur.

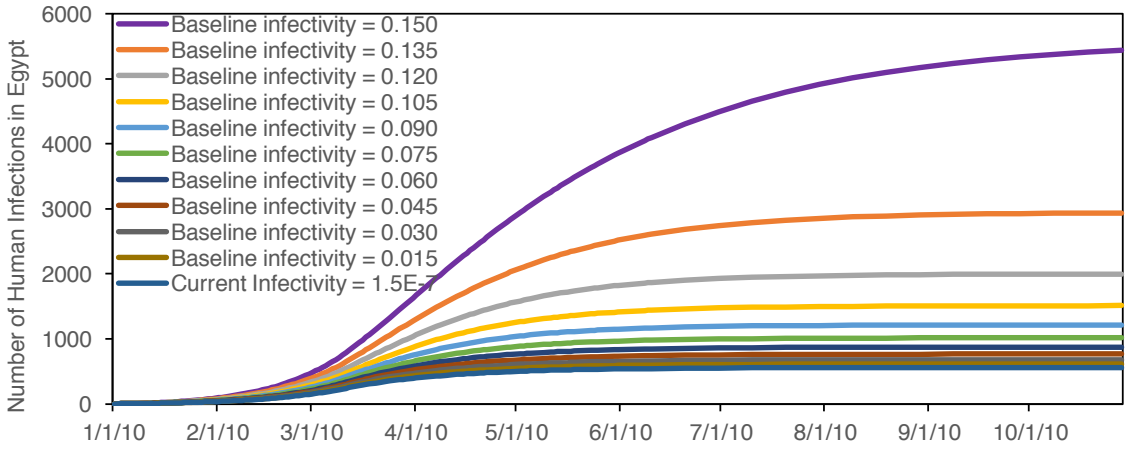


Figure 1.13: Avian influenza: sensitivity analysis of baseline infection rate between humans in Egypt.

United States. In this scenario, besides the factors discussed in the previous section, we also explore the impact of transportation and migration of infectious birds on disease trend of an avian influenza outbreak. For brevity, we only consider infections among birds.

To model the change of bird population in each state due to commercial transportation or migration, we introduce a transportation matrix $\boldsymbol{\Omega}_{N\times N} = \langle \omega_{ab} \rangle|_{1\le a,b\le N}$, where ω_{ab} represents the percentage of birds transported from state a to state b in a unit time, and N is the total number of states considered in this setup. Each state is assigned a unique instance of Equation 1.10b with its own initial bird populations, and the transportation matrix $\boldsymbol{\Omega}$ is used to connect all the states. More rigorously, let $\boldsymbol{Y}_{m\times N} = [\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_N]$ be the collection of stage vectors of all states, where m is the number of total compartments for birds in Equation 1.10b. Then the ordinary differential equation with the transportation and migration of birds can be written as

$$\begin{aligned}\boldsymbol{Y}' =& (\boldsymbol{D}_B - \boldsymbol{I})\text{diag}(\boldsymbol{\mu}_B)\boldsymbol{Y} + \sum_{k=1}^{N} \text{diag}((\boldsymbol{D}_B + \boldsymbol{I})\text{diag}(\boldsymbol{\beta}_{B\to B})\boldsymbol{y}_k)\boldsymbol{C}_{B\to B}\boldsymbol{y}_k\boldsymbol{e}_k^\intercal \\ &+ \lambda_B(\boldsymbol{Y}) + \boldsymbol{Y}(\boldsymbol{\Omega} - \text{diag}(\boldsymbol{\Omega}\cdot\mathbf{1}))\end{aligned}$$

where $\boldsymbol{e}_k$ is the k-th standard basis vector. In this scenario, we assume that the initial infection starts in Ohio with 30,000 infectious birds. If no restrictions are placed on the transportation of live poultry and no interventions are implemented, the spread of the virus is shown in Figure 1.14, where the upper panel shows the real-time infections, and the lower panel shows the cumulative infections. By day 100 of the outbreak, the virus has already spread to neighboring states and most of the northeast region. By day 300, the virus has spread to the midwest and southern regions, and the trend of the outbreak is still increasing, as shown in the light green solid curve in Figure 1.15. The cumulative number of infections in birds in the entire United States will reach 1.8 million by day 300, and some states with high poultry production will be severely affected, causing potential dangers to the health of poultry farm workers as well as residents.

We will examine two intervention strategies for this scenario: 1) reducing the probability of transporting infectious birds to other states by enforcing strict regulations and

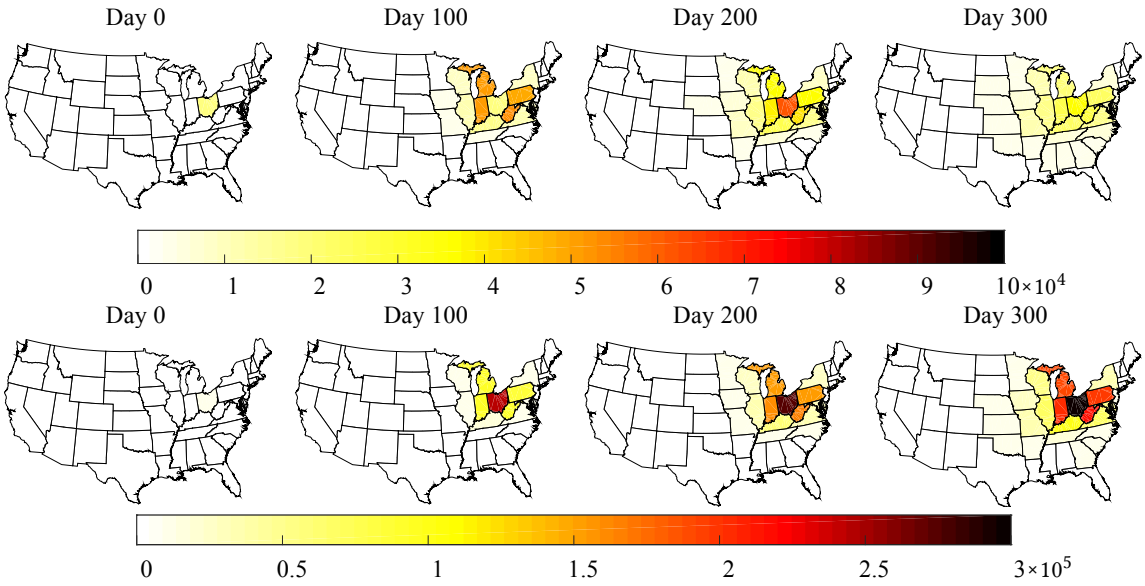


Figure 1.14: Avian influenza: spread in US after the initial outbreak.

examinations on the transportation of live poultry and/or poultry products; 2) reducing the population of infectious birds by slaughtering infectious poultry and wild birds. Both strategies are assumed to be implemented one month after the initial outbreak.

Figure 1.15(a) shows the effect of reducing transportation of infectious live poultry and controlling the migration of infectious wild birds. Without any intervention, the outbreak will start to show signs of containment by day 700, and the total number of infectious birds at containment will be near 2.6 million. If a 20% reduction of infectious transportation is achieved, the containment will start on day 200, and the total infections in bird population at containment will be 1.1 million, a 58% reduction. If this strategy is implemented at 80% level, the outbreak will start to be contained by day 120 (almost one and a half years earlier compared to no intervention), and the total infection in bird population at containment will be 0.6 million, a 77% reduction. This result demonstrates that by imposing strict regulations and examinations on live poultry transportation, the containment will be reached earlier, and the total number of infected birds will be reduced significantly. Another ad-

vantage of this strategy is that it can confine the outbreak within a small region so that it is easier to implement other types of interventions.

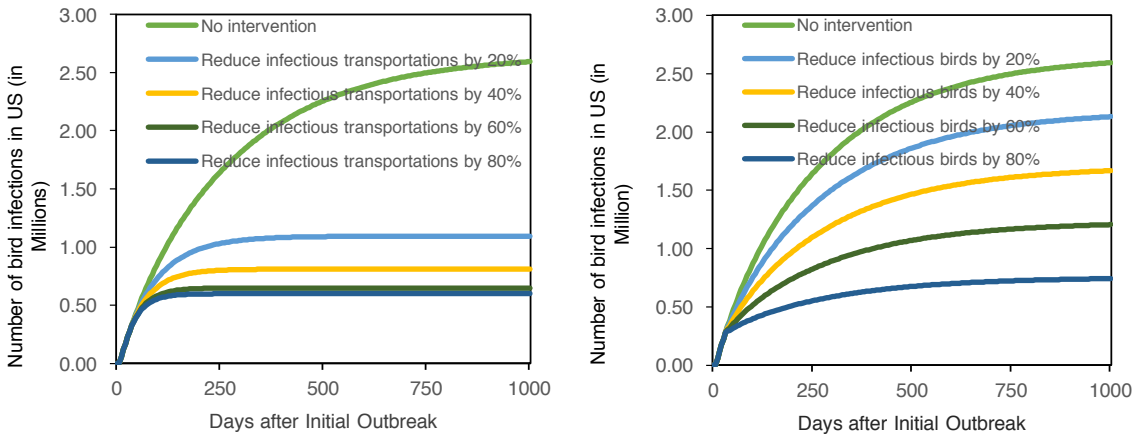


Figure 1.15: Avian influenza: bird infections in US under two containment strategies.

Figure 1.15(b) shows the effect of slaughtering infectious birds one month after the initial outbreak. Although slaughtering infectious birds alone is not as effective as reducing transportation and does not change the disease trend significantly, it does reduce the number of infections at containment. For example, when 40% of infectious birds are slaughtered after one month of the initial outbreak, the cumulative number of infections at containment will be 1.67 million, a 36% reduction compared to no intervention. The number of infections at containment will be 0.74 million, a 72% reduction, when 80% of infectious birds are slaughtered. However, this strategy has both environmental and economic concerns, and the disposal of carcasses of slaughtered birds remains a difficult problem to solve30. Combination of containment strategies will be effective and promising, as shown in Figure 1.16. Implementing both strategies at 20% level will reduce the number of infections by containment to 0.95 million, a 63% reduction. When both strategies are implemented at 30% level, the cumulative number of infected birds will be 0.75 million, a 71% reduction.

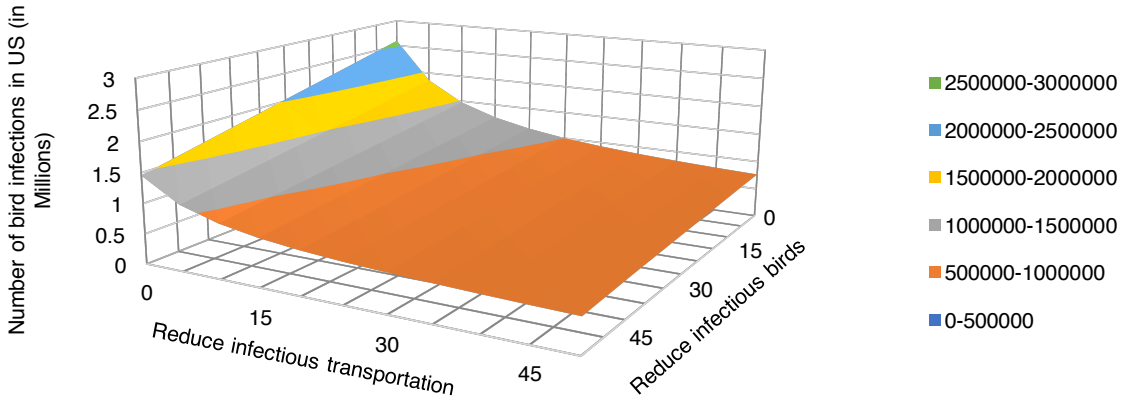


Figure 1.16: Avian influenza: effect of combination strategies in US.

Discussion

This work signifies the idea of "digital disease surveillance" in preparing for and responding to pandemic outbreaks. We develop a general-purpose modeling framework for digital surveillance that couples compartmental models (with separate groups of compartments for human and carrier populations) with human behavior and operations processes. We demonstrate its usage for analyzing avian influenza. The modeling framework is capable of tracking the development (and interplay) of epidemics in both groups. It also incorporates the natural population growth and effect of migration and transportation into consideration. We analyze the development of the avian influenza outbreaks and investigate parameters that have significant impacts on epidemic severity. We study the effect of average contact rate between birds and the average time that birds stay infectious. The analyses are performed using the real data from an outbreak in 2010 in Egypt. Our analyses show that the trend of avian influenza outbreaks is sensitive to both factors, and accurate parameter estimation of the biological properties of avian influenza viruses is crucial to the performance of the model.

We explore three containment strategies for the outbreak data from Egypt to mitigate and minimize infections among humans: reducing human contact rate to birds, slaughtering

infectious birds/live poultries, and quarantining infectious humans. The first two strategies show similar effectiveness in containing the outbreak and reducing the number of infections in humans, while there is marginal effect in the quarantining infectious humans. This shows that the key factor in containing an avian influenza outbreak lies in controlling the infections in birds and reducing the chances of human contacting infectious birds. The human-to-human infections effect of avian influenza is low; however, it is possible that mutations may lead to more significant human-human infections.

Analyses are also carried out using the poultry farm and production data in the United States. Two containment strategies are contrasted: 1) reducing transportation or migration of infectious birds, and 2) slaughtering the infectious bird population. Both strategies are effective, with the former confining the outbreak within a localized region for more manageable containment operations. These findings indicate that regardless of the objectives, the outbreak in the bird population should be tackled first, and the containment strategy should be determined based on environmental and demographic information, as well as transportation/movement and bird migration patterns.

The model provides a decision support framework for policymakers to estimate the potential risk of an outbreak and the cost-effectiveness for different containment strategies. Combination strategies turn out to be practical and cost-effective in achieving disease-free equilibrium. Healthcare departments should select a strategy portfolio in accordance with the local environment, demographics, and the composition of the bird population. To further expedite the containment, the public should be educated and informed of the knowledge regarding avian influenza, as human behaviors have a profound impact on the virus transmission.

Informatics and social media play an important role in tracking disease spread and sup-

porting epidemiology modeling and containing disease outbreaks and have already demonstrated their capability in helping prevent influenza outbreaks [211]. The most challenging part is that how classical mathematical models fit and perform in this digital era. Our proposed modeling framework is also highly sustainable with the support of modern information technologies, which provide valuable inputs to our model. With the continuous data input, the model can adjust the parameter estimations to better predict the potential outcomes of pandemics, reflect changes in the environment, human behavior, and public policies as the outbreak develops and provide meaningful insights to the containment strategies in real-time to better facilitate the decision-making process. It can also be easily cast to accommodate newer disease propagation structures, additional stages introduced by intervention or hospitalization, and changes in the pattern of disease transmission, making it sustainable both theoretically and practically.

There are limitations to our study. Although the modeling framework we proposed is flexible and capable of modeling complicated scenarios, we used a relatively simple setup in the examples with homogeneous bird and human populations without any treatment or vaccination effects. Different bird species have various infectivity regarding avian influenza, and the contact rate for birds differs between poultry farm workers and general civilians. Antiviral drugs and vaccines against avian influenza also exist and could be added to the model. These components require more parameters to estimate and may possibly lead to over-fitting. However, with the availability of such data, these effects can all be included in our modeling framework with additional compartments in both bird and human population.

CHAPTER 2
DISEASE PROPAGATION MODELS FOR MEDICAL OPERATIONS AND COUNTERMEASURES

In this chapter, we extend the modeling framework proposed in Chapter 1 to incorporate the impact of medical operations and countermeasures on the trend of infectious disease outbreaks in two different ways: 1) optimizing the distribution of limited medical resource to achieve the best containment result; 2) modeling the disease propagation and transition of stages in the population inside and outside medical facilities to understand role medical operations and countermeasures plays during pandemics. Then we discuss two applications of the proposed models to assist decision-making in the Ebola virus outbreak in West Africa and the optimal prioritized vaccination strategy for a smallpox outbreak.

2.1 Literature Review

Decision-making during infectious disease outbreaks is crucial to achieving early epidemic containment, minimizing the total number of infections, and optimally allocating medical resources. Infectious disease models can be used to understand the effects of containment efforts and provide guidance for implementing interventions during outbreaks. Compartmental models provide accurate forecasts for the potential development of outbreaks and incorporate complex system dynamics of infectious disease propagations. They have been used to assist developing intervention strategies and evaluating the performance and efficiency of various medical operations and countermeasures during epidemics before implementation to help decision-makers identify the most cost-effective strategy and unveil potential influence of different types of interventions. Multiple studies have introduced human interventions and behavior changes to the existing compartmental models to investigate their effect.

An intuitive way to extend the current compartmental model is to include multiple additional stages introduced by human interventions combined with healthcare system modeling. Legrand et al. studied the dynamics of Ebola pandemics using data from two different pandemics, quantified the transmission rates with various scenarios including illness in the community, hospitalization, and traditional burial, and ran simulations of these scenarios to explore the impact of control interventions on a potential epidemic [145]. Rivers et al. used data from Liberia and Sierra Leone in the 2014 West Africa Ebola outbreak to parameterize a mathematical model for forecasting the efficacy of different intervention methods including increased contact tracing, improved infection control practices, and the use of a hypothetical pharmaceutical intervention [185]. Lee et al. studied the effect of vaccination and special treatment during an influenza outbreak with a limited supply of vaccines by extending the existing 6-stage compartmental model [143]. Fukuda et al. assessed the effect of voluntary vaccination and its impact on a network structure, combined the human intervention and behavior changes during disease outbreaks, and concluded that the public information has a positive impact on containment given the vaccination cost is small [92].

Among all intervention strategies, hospital-based interventions and community-based interventions are widely used, and mathematical models have been developed to evaluate the impact of these two types of intervention strategies during pandemic outbreaks. Riley et al. studied the transmission dynamics of the severe acute respiratory syndrome (SARS) in Hong Kong and addressed the importance of shortening the delay between the onset of symptoms to hospitalization as well as improving the isolation effect inside hospitals [184]. Nkoghe et al. conducted a case study of a regional Ebola virus outbreak in Republic of Congo in 2005 and discussed the impact of isolation ward and proper training of healthcare workers within hospitals on the effect of reducing the number of new infections and preventing in-hospital transmission [171]. Pandey et al. developed a stochastic model for Ebola outbreak in West Africa and evaluated the virus transmissions within

and between different communities with data from Liberia. They concluded that isolation, contact-tracing with quarantine, and sanitary funerals must be prioritized to control the outbreak. They also found that an insufficient amount of beds and low admission rates to hospitals reduce the effectiveness of hospital-based interventions [174]. Milne et al. proposed a small community model for the spread of infectious diseases and used the household structure and exact individual-level contact patterns to model the spread of influenza. They evaluated four community-based interventions including school closure and increased case isolation and concluded that multiple social distancing strategies need to be implemented early and continuously for optimal transmission interruption [163].

With the availability of modern transportation, global pandemic outbreaks are more probable to occur and more difficult to mitigate [53]. Global-level intervention strategies have been proposed and studied, aiming to provide quantitative analysis to support the decision-making processes. Balcan et al. proposed the global epidemic and mobility model which combines sociodemographic and population mobility data and a stochastic disease propagation model with a spatial structure to predict the spread of infectious diseases at the global scale and discussed how intervention strategies can be incorporated into this modeling framework to mitigate and contain epidemics [18]. Gomes et al. studied the risk of international spreading the 2014 Ebola outbreak in West Africa based on the global epidemic and mobility model, using agent-based simulation to predict the spread of Ebola virus worldwide. They showed that local interventions to mitigate regional outbreaks and traveling interventions to reduce case importation are both effective policies to prevent global Ebola pandemic [98]. Similarly, Poletto et al. assessed the impact of the travel ban to affected areas of the 2014 Ebola outbreak and used numerical simulations to evaluate the results. They concluded that travel ban alone only delays the risk of outbreaks extending to other countries, and additional intervention strategies need to be implemented to prevent further geographical spread [179]. McCloskey et al. addressed the importance of effective

collaboration between countries within a unified framework to achieve rapid detection of pandemics and start public health responses immediately after the initial outbreak [157].

2.2 Resource Allocation for Containing Disease Outbreaks

During infectious disease outbreaks, the supply of medical resources and the availability of healthcare workers are always not enough to implement the containment strategies at desired levels. Therefore, optimal allocation of limited resources is crucial for achieving early containment and minimizing total infections. In this section, we discuss a resource allocation problem based on the general-purpose disease modeling framework proposed in Chapter 1 and provide a numerical method to solve the problem.

2.2.1 Additional Stages for Intervention Modeling

Due to the human interventions during the disease outbreaks, new stages that are not natural to the disease propagation will be introduced to our existing modeling framework. An example of stages of human intervention is hospitalization. Individuals in the hospitalization stage will have a higher recovery rate and a lower mortality rate due to treatment, and a lower contact rate to the populations in other stages. However, the availability of hospitalization is constrained by the number of beds in the healthcare facilities, the number of healthcare workers and doctors, and the availability of necessary medical resources. Other examples of stages introduced by human intervention include vaccination, where susceptible individuals can become vaccinated and have a lower probability of getting infected, constrained by the availability of vaccines and the processing rate at vaccine clinics; and quarantine, where infectious individuals will have a lower contact rate to the healthy groups. Since infectious individuals can perform self-quarantine at home, the availability is usually not constrained.

If we focus on the macroscopic aspect of the disease propagation, i.e., ignore the de-

tailed operations within the stages introduced due to human interventions and treat them as black boxes, the modeling framework proposed in Chapter 1 will be sufficient to incorporate these new stages. Consider the SEIR model with an additional stage for hospitalization H shown in Figure 2.1.

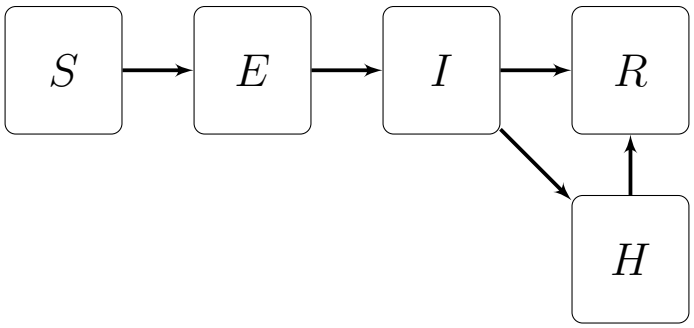


Figure 2.1: Demonstration of an SEIR model with hospitalization.

Assume that individuals in stage H are quarantined, then using the categorization of stages, $H \in \Phi_U$ and $H \in \Phi_A$. Assume the mean time for an individual from becoming infectious to being hospitalized is $1/\mu_T$ and arrange the stages as $\{S, E, I, R, H\}$, then the model matrices of the modeling framework becomes

$$\boldsymbol{D} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & \mu_I/(\mu_I+\mu_T) & 0 & 1 \\ 0 & 0 & \mu_T/(\mu_I+\mu_T) & 0 & 0 \end{bmatrix}, \boldsymbol{C} = \begin{bmatrix} 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

If we further assume the mean duration for hospitalization is $1/\mu_H$, then the rest of model parameters are

$$\boldsymbol{\mu} = (0, \mu_E, \mu_I + \mu_T, 0, \mu_H)^\intercal, \boldsymbol{\beta} = (\beta, 0, 0, 0, 0)^\intercal$$

Plugging these modeling matrices into the general-purpose modeling framework for contact-based diseases will yield the system of ordinary differential equations that capture human interventions. On the other hand, if the individuals in stage H are not fully quarantined, but contact populations in other stages at a reduced rate, then we may modify the disease contagious matrix $\boldsymbol{C}$ for this scenario. In this case, $H \in \Phi_C$. Let $\boldsymbol{\eta} = \langle \eta_\phi \rangle|_{\phi \in \Phi}$ denote a vector of contact rate modification factor for each compartment $\phi \in \Phi$. For stage ϕ that does not correspond to human interventions, $\eta_\phi = 1$; for stage ϕ that correspond to the contagious stage but with a diminished contact rate, $\eta_\phi < 1$. Different from the previous example, since now $H \in \Phi_C$, the disease contagious matrix becomes

$$\boldsymbol{C} = \begin{bmatrix} 0 & 0 & -1 & 0 & -1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Then we can substitute the matrix $\boldsymbol{C}$ with $\boldsymbol{C}\text{diag}(\boldsymbol{\eta})$ in the modeling framework to reflect the change of contact rate due to human interventions.

2.2.2 Modeling Framework with Interventions

In our modeling framework, human intervention may impact the spread of infectious diseases in two ways: affecting the development of the disease by hospitalization, treatment, or vaccination, and affecting the generation of new infections by quarantine and reducing contact rate. Start with the case of contact-based diseases. Mathematically, since the contagious structure of the infection is inherent to the disease itself, the disease contagious matrix $\boldsymbol{C}$ will not be changed by human interventions. The first type of intervention changes the mean transition rate at the active stages, i.e., $\boldsymbol{\mu}$. By altering $\boldsymbol{\mu}$, the disease development matrix $\boldsymbol{D}$ may also be changed; the second type of intervention changes the baseline infection

rate $\boldsymbol{\beta}$.

Let vector $\boldsymbol{r}(t)$ denote the availability of multiple medical resources at time t, and the value of $\boldsymbol{r}(t)$ has impact on both $\boldsymbol{\mu}$ and $\boldsymbol{\beta}$. Consider two functions $f(\boldsymbol{\mu}, \boldsymbol{r}(t))$ and $g(\boldsymbol{\beta}, \boldsymbol{r}(t))$ that output the updated value of mean transition rate of active stages $\boldsymbol{\mu}$ and baseline infection rate of vulnerable stages $\boldsymbol{\beta}$ at time t with the available medical resource $\boldsymbol{r}(t)$. As the disease development matrix $\boldsymbol{D}$ may rely on the value of $\boldsymbol{\mu}$ but not directly controlled by the available medical resource $\boldsymbol{r}$, the updated disease development matrix can be written as $\boldsymbol{D}(f(\boldsymbol{\mu}, \boldsymbol{r}(t)))$. Then we may substitute the constants $\boldsymbol{\mu}$, $\boldsymbol{\beta}$ and $\boldsymbol{D}$ in the original modeling framework with the time-variant parameters described above to model the effect of different ways of human interventions.

Since interventions will change the level of medical resource inventory and there may be possible resupplies and other consumptions of the medical resources, the derivative of $\boldsymbol{r}(t)$ should also be fully explored to correctly model the system dynamics with interventions. The consumption rate of medical resources is fully described by the current change rate of population among all stages $\boldsymbol{y}'(t)$ and the current level of medical resources $\boldsymbol{r}(t)$ in the form of $\xi(\boldsymbol{y}'(t), \boldsymbol{r}(t))$. On the other hand, assume the rate of resupply and other consumptions of medical resources is $\boldsymbol{\pi}(t)$, then the change rate of medical resources can be written as $\boldsymbol{r}'(t) = \xi(\boldsymbol{y}'(t), \boldsymbol{r}(t)) + \boldsymbol{\pi}(t)$. For simplicity, we will omit the time variable t in the following discussions. Combining the updated expression for $\boldsymbol{y}'$ and the equation for $\boldsymbol{r}'$, the general-purpose modeling framework for contact-based infectious diseases with human intervention can be written as

$$\begin{aligned}\boldsymbol{y}' = &(\boldsymbol{D}(f(\boldsymbol{\mu}, \boldsymbol{r})) - \boldsymbol{I})\text{diag}(f(\boldsymbol{\mu}, \boldsymbol{r}))\boldsymbol{y} \\ &+ \text{diag}((\boldsymbol{D}(f(\boldsymbol{\mu}, \boldsymbol{r})) + \boldsymbol{I})\text{diag}(g(\boldsymbol{\beta}, \boldsymbol{r})\boldsymbol{y}))\boldsymbol{C}\text{diag}(\boldsymbol{\eta})\boldsymbol{y} \qquad (2.1a)\end{aligned}$$

$$\boldsymbol{r}' = \xi(\boldsymbol{y}', \boldsymbol{r}) + \boldsymbol{\pi} \qquad (2.1b)$$

Consider the SEIR model with hospitalization discussed in the previous section, but now assume that the hospitalization stage has a capacity M. If the number of hospitalized individuals reaches M, no more infectious individuals will be treated. In this example, there is a single medical resource $r = \max\{M - H, 0\}$, which is the vacancies in hospitals, and it has only impact on the vector $\boldsymbol{\mu}$. Therefore, $g(\boldsymbol{\beta}, \boldsymbol{r}) = \boldsymbol{\beta}$. To understand how the vacancies in hospitals impact the vector $\boldsymbol{\mu}$, notice that the transition rate from infectious stage to hospitalization stage is μ_T if $r > 0$, and is 0 when $r = 0$. Therefore, the transition rate from infectious stage to hospitalized stage is a variable $\mu_T \mathcal{I}\{r > 0\}$, where $\mathcal{I}$ is the indicator function. Therefore,

$$f(\boldsymbol{\mu}, r) = (0, \mu_E, \mu_I + \mu_T \mathcal{I}\{r > 0\}, 0, \mu_H)^\intercal$$

If we assume that the capacity of hospitals remains constant throughout the outbreak, then $\pi(t) = 0$ for all $t > 0$. Thus the consumption and resupply of the medical resource r is solely determined by the change in the size of hospitalized individuals under this assumption, i.e., $\xi(\boldsymbol{y}', r) = -\boldsymbol{e}_H^\intercal \boldsymbol{y}' = -H'(t)$. Therefore, $r'(t) = -H'(t)$. Then we have fully characterized the spread of infectious disease in an SEIR model with capacitated hospitalization using the models described in Equation 2.1. Notice that this modeling framework can be easily extended to the vector-borne diseases by specifying the effect of medical resources $f(\boldsymbol{\mu}, \boldsymbol{r})$ and $g(\boldsymbol{\beta}, \boldsymbol{r})$ on both the human and vector populations.

2.2.3 Optimal Resource Allocation under the Modeling Framework

With human interventions and the use of medical resources properly incorporated in the modeling framework for infectious disease outbreaks, the resource allocation problem can be formulated as an optimization problem in two ways: minimizing the medical resources needed throughout the outbreak with constraints on cumulative infections and/or containment time; and minimizing total infections and/or containment time with constraints on

the availability of medical resources throughout the outbreak. Let $u(t)$ denote some metrics related to the population vector $\boldsymbol{y}$ at time t which we are interested in to evaluate the performance of different strategies of interventions. For example, if we are interested in minimizing the total cumulative infections up to time t_0, then

$$u(t_0) = -\int_0^{t_0} \boldsymbol{e}_V^\intercal \boldsymbol{y}'(t) \mathrm{d}t$$

where $\boldsymbol{e}_V$ is a $|\Phi|$-dimensional vector with value 1 for vulnerable stages and 0 otherwise. On the other hand, if we are interested in minimizing the time needed to reach outbreak containment, we can define the metric variable at time t_0 as $u(t_0) = \mathcal{I}\{\boldsymbol{e}_V^\intercal \boldsymbol{y}'(t) < \epsilon\}$, where ϵ is a threshold to determine if containment is achieved.

Similar to the function $u(t)$ that provides a metric to evaluate the extent of disease outbreaks, we can define a function $v(t)$ that evaluates the value of the investment of medical resources at time t during the outbreak. One intuitive example is the total quantity of the medical resources. The medical resource minimization problem with constraints on the performance can be formulated as

$$\begin{aligned} &\min \max_{t \in T} v(t) \text{ or } \min \int_{t \in T} v(t) \mathrm{d}t \\ &\text{s.t. } u(t) \leq u_{\text{goal}}(t) \quad \forall t \in T \end{aligned}$$

where the first type of objective function is suitable to model the usage of medical resources that are not consumable, like vacancies in hospitals or the availability of healthcare workers, and the second type of objective function can be used to model the usage of consumable resources including medications, vaccines, etc. The constraint ensures that for any time in the horizon of the outbreak T, the value of the evaluation metric $u(t)$ reflecting the severity of outbreak should be bounded by a pre-defined value $u_{\text{goal}}(t)$. Similarly, the total

infections minimization problem with constraints on the availability of medical resources can be formulated as

$$\begin{aligned} &\min\ u(t_0) \\ &\text{s.t. } v(t) \leq v_{\text{available}}(t) \quad \forall t \leq t_0 \end{aligned}$$

where the objective function is the metric $u(t)$ evaluated at time t_0, i.e., the severity of the outbreak at time t_0. The constraint ensures that during the time frame $[0, t_0]$, the usage of medical resources should not exceed the availability of the resource $v_{\text{available}}(t)$ at any time.

To solve the resource allocation optimization problem, notice that evaluating a single value of the objective function will require solving the entire system of ordinary differential equations until equilibrium is reached. Therefore, we use the following line search algorithm to numerically solve the optimization problem:

1. First solve the system of ordinary differential equations until equilibrium is reached and evaluate the value of objective function;

2. Introduce an increment of medical resources $\Delta \boldsymbol{r}$ to the system and solve the updated ordinary differential equations until equilibrium is reached;

3. Repeat step 2 until the constraint is satisfied for the first type of optimization problem, or all medical resources are used up (i.e., the constraint is violated) in the second type of optimization problem.

This algorithm is straightforward: we gradually introduce the interventions to the system and observe the change of the trend of the outbreak caused by the intervention; repeat this process until the constraints are satisfied or the resources are used up. This is also an online algorithm, as the best solution observed so far can always be obtained when the algorithm is interrupted. Since solving a system of ordinary differential equations to

equilibrium is still computationally expensive, the following two strategies can be used to accelerate the solving procedure:

* When the value of the objective function improves significantly (not saturated), a larger step size of the increment of the medical resource can be used in each iteration. When the improvement of the objective function starts to slow down, the step size can be tuned down. This is similar to adjusting the learning rate in many machine learning algorithms.

* The commonly used algorithm to numerically solve ordinary differential equations is the fourth order Runge-Kutta method by averaging 4 increments, each one requiring calculating $\boldsymbol{y}'$ from scratch. However, infectious disease outbreaks either contain within a certain duration or end up with infecting every individual in the environment in practice. Therefore, the trend of the spread of infectious disease always follows the "S" curve, i.e., the outbreak would go through an "initial growth–fast growth–containment" cycle. The behavior of the system is highly regularized and we may assume that the rate of change within a sufficiently small time interval is constant. With this observation, we can use the first order approximation (Euler method) and truncate the Taylor series at order 2 to solve the ordinary differential equations and reduce the amount of calculation by ignoring the higher order terms. Our simulation shows that the predictions of fourth order and first order approximation are almost the same in most applications of disease propagation modeling.

With properly defined medical resources $\boldsymbol{r}$, the evaluation metrics for the severity of outbreak $u(t)$, and the availability of medical resource $v(t)$, the resource allocation problem can be defined based on the general-purpose disease modeling framework and the optimal intervention strategies can be obtained through solving the resource allocation problem.

2.3 Modeling Medical Operations and Countermeasures

In this section, we discuss how to model the medical operations and countermeasures during infectious disease outbreaks in a more microscopic aspect, instead of treating it as a black box in the previous section. This modeling technique is used to determine if the performance and the operation efficiency of healthcare workers are the bottlenecks of achieving fast containment, and can assist decision-makers to adjust the local operation strategies accordingly for a better containment result.

2.3.1 POD Layout and Separation of Environment

In the following discussion, we assume that all the medical operations and countermeasures happen inside point-of-dispensing sites (PODs). To clearly define what happens inside a POD, the layout of the POD should be taken into consideration in the models. Figure 2.2 demonstrates a sample layout design for a POD: clients will arrive at the POD and visit the triage station first. If they need further medical or mental assistance, they will be sent to the corresponding station and the healthcare workers at that station will determine if they are suitable to receive drugs or vaccines; otherwise, they will be sent directly to the dispensing station.

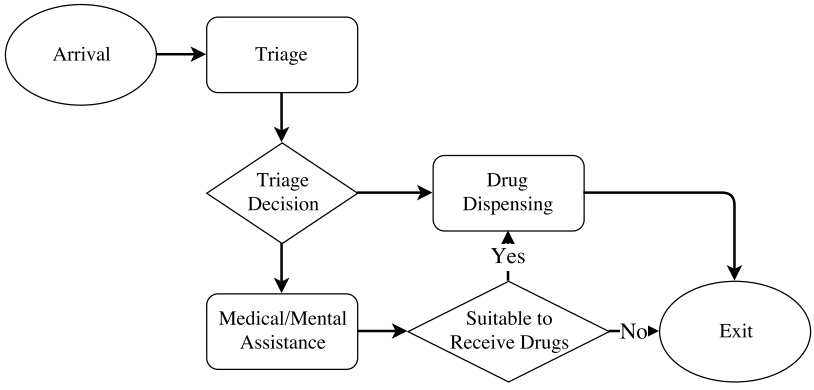


Figure 2.2: A sample POD layout design.

Similar to the discussion in the previous section, to model the effect of human interven-

tions inside the PODs, new stages need to be introduced to the original modeling framework. The new stages can still be described using the categorization method discussed earlier, but they do not belong to the natural process of disease propagation and cannot be modeled using the disease transition matrix $\boldsymbol{D}$; instead, such transition happens if and only if when the proper intervention is performed on the corresponding population. Therefore, $d_{ab} = 0$ if either stage a or stage b belongs to the compartments corresponding to human interventions. In addition, we separate the environment in which the disease propagation happens into two parts to depict the effect of PODs in the process of the disease outbreak: intra-POD and outer-POD disease propagation. This separation also takes the different population densities in the two environments into consideration, thus can correctly model the increased contact rate due to the high population density inside the PODs.

The outer-POD disease propagation can be described using the same modeling framework discussed in Chapter 1. However, the intra-POD disease propagation has two additional layers of complexity: the status of individuals inside the PODs can change due to the natural propagation of infectious diseases, the service provided inside the PODs, and the movement of individuals according to the POD layout. To properly capture this property of intra-POD disease propagation, assume the considered POD has K different stations, each of which provides certain service to the arriving clients. We assign a population vector $\boldsymbol{y}_i, i \in \{1, 2, \ldots, K\}$ to each of the K stations and let $\boldsymbol{y}_0$ denote the population vector outside the PODs. Therefore, besides the disease propagation model discussed in Chapter 1, two additional relationships need to be established between $\boldsymbol{y}_i, i = \{1, 2, \ldots, K\}$ to model the change of population due to the medical operations provided by the PODs and the movement of individuals inside the PODs.

Assume that station i is manned with n_i healthcare workers, each of them provides the same service at an identical nominal rate λ_i. Therefore, the actual service rate (departure

rate) at station i for populations in stage ϕ is

$$h_i(\phi) = \min\{n_i, ||\boldsymbol{y}_i||_1\}\frac{\lambda_i \phi_i}{||\boldsymbol{y}_i||_1}$$

where ϕ_i is the number of individuals in stage ϕ at station i. Intuitively, the overall service rate at station i is $\min\{n_i, ||\boldsymbol{y}_i||_1\}\lambda_i$. Since individuals from multiple disease groups are served at station i, this overall service rate is distributed across all the disease groups according to the relative sizes of each disease group, hence the factor $\phi_i/||\boldsymbol{y}_i||_1$ is used to adjust the service rate for each disease stage.

After receiving service at certain stages, clients can be transited to other stages. For example, a susceptible individual will become vaccinated with a certain probability after the service of vaccination stages. Since this type of stage transitions is not natural to the disease propagation itself, it is not characterized using disease transition matrix $\boldsymbol{D}$. Instead, we use a *stage-jump matrix* $\boldsymbol{J}_i = \langle j^i_{ab} \rangle|_{a\in\Phi, b\in\Phi}$, where j^i_{ab} is the probability of transiting individuals from stage b to stage a at station i, to summarize the change of stages due to the service performed at station i. If the service at station i does not change the status of individuals in disease stage ϕ, then $j^i_{\phi\phi} = 1$. Therefore, the summation of each column of the stage-jump matrix $\boldsymbol{J}_i$ is always 1. We use a matrix $\boldsymbol{T}$ to compactly summarize the output rate at each station:

$$\boldsymbol{T} = [\boldsymbol{J}_1 h_1(\Phi), \boldsymbol{J}_2 h_2(\Phi), \ldots, \boldsymbol{J}_K h_K(\Phi)]$$

where the i-th column of the matrix $\boldsymbol{T}$ indicates the rates at which the individuals from different disease stages leaving from station i.

Besides the transitions between disease stages, clients move inside the PODs according to its layout and design. To model the change of population due to the movement

inside PODs, we define a *station-jump matrix* for each station inside the POD: $\boldsymbol{Q}_i = \langle q^i_{ab} \rangle|_{0 \le a \le K, b \in \Phi}$ such that q^i_{ab} is the probability of clients in disease stage b moving to station i after visiting station a. For completeness, we use $\boldsymbol{Q}_0$ to represent the matrix of the probability of clients in all disease stages moving from station a to the outer-POD environment. This establishes the connection between intra-POD and outer-POD environments at the exit of PODs. To build such connection at the entrance of PODs, let $\psi_i, i = \{1, 2, \ldots, K\}$ be an indicator variable of whether station i is the first station to visit inside the POD. Additionally, let $\boldsymbol{\lambda} = \langle \lambda_\phi \rangle|_{\phi \in \Phi}$ be the baseline arrival rate of individuals in each stage outside the POD. To better reflect the realistic and typical emergency response in case of a pandemic outbreak, we assume that the arrival rate of the PODs is a time-variant parameter instead of a constant throughout the horizon of the disease outbreak. Therefore, the actual arrival rate at PODs is the baseline arrival rate $\boldsymbol{\lambda}$ adjusted by the total populations outside of POD:

$$\boldsymbol{\lambda}_{\text{adjusted}} = \frac{\text{diag}(\boldsymbol{\lambda})\boldsymbol{y}_0}{||\boldsymbol{y}_0||_1}$$

2.3.2 Modeling Framework for Medical Operations and Countermeasures

With the model parameters for medical operations and countermeasures defined in section 2.3.1, the modeling framework for this application can now be derived. Consider the population dynamics at station i inside the PODs. The change of populations can be decomposed into four parts: change due to natural disease propagation, change due to clients arrival from the outer-POD environment, change due to the processing rate at station i, and change due to the arrivals from upstream stations. We use $\boldsymbol{y}'_{i,1}, \boldsymbol{y}'_{i,2}, \boldsymbol{y}'_{i,3}$ and $\boldsymbol{y}'_{i,4}$ to denote the change of population due to these four reasons.

Obviously, $\boldsymbol{y}'_{i,1}$ does not involve in the operations inside the PODs and should remain the same as the modeling framework discussed in Chapter 1; $\boldsymbol{y}'_{i,2}$ only depends on whether the i-th station is the first station to visit in the POD; and $\boldsymbol{y}'_{i,3}$ is solely determined by the

processing rate at station i. Use the contact-based disease as an example, we have

$$\begin{aligned}
\boldsymbol{y}'_{i,1} &= (\boldsymbol{D} - \boldsymbol{I})\text{diag}(\boldsymbol{\mu})\boldsymbol{y}_i + \text{diag}((\boldsymbol{D} + \boldsymbol{I})\text{diag}(\boldsymbol{\beta})\boldsymbol{y}_i)\boldsymbol{C}\boldsymbol{y}_i \\
\boldsymbol{y}'_{i,2} &= \psi_i \frac{\text{diag}(\boldsymbol{\lambda})\boldsymbol{y}_0}{||\boldsymbol{y}_0||_1} \\
\boldsymbol{y}'_{i,3} &= -h_i(\Phi)
\end{aligned}$$

To derive the expression of $\boldsymbol{y}'_{i,4}$, consider the row that corresponds to stage ϕ in matrix $\boldsymbol{T}$: $\boldsymbol{t}_{\phi\cdot} = [t_{\phi 1}, t_{\phi 2}, \ldots, t_{\phi K}]$. Each element in this row vector shows the departure rate of individuals in stage ϕ from each station inside the POD. Then consider the column corresponds to stage ϕ in matrix $\boldsymbol{Q}_i$: $\boldsymbol{q}_{\cdot\phi} = [q_{1\phi}, q_{2\phi}, \ldots, q_{K\phi}]^\intercal$, where each element in this column vector is the percentage of individuals in stage ϕ from each station that enters station i. Therefore, the total population in stage ϕ that enters station i from all other stations can be expressed as $\boldsymbol{t}_{\phi\cdot}\boldsymbol{q}_{\cdot\phi}$. Written compactly in matrix representation,

$$\boldsymbol{y}'_{i,4} = \text{diag}(\boldsymbol{T}\boldsymbol{Q}_i)\boldsymbol{1}$$

Then the system of ordinary differential equations for the population dynamics at station i is

$$\begin{aligned}
\boldsymbol{y}'_i = {} & (\boldsymbol{D} - \boldsymbol{I})\text{diag}(\boldsymbol{\mu})\boldsymbol{y}_i + \text{diag}((\boldsymbol{D} + \boldsymbol{I})\text{diag}(\boldsymbol{\beta})\boldsymbol{y}_i)\boldsymbol{C}\boldsymbol{y}_i \\
& + \psi_i \frac{\text{diag}(\boldsymbol{\lambda})\boldsymbol{y}_0}{||\boldsymbol{y}_0||_1} - h_i(\Phi) + \text{diag}(\boldsymbol{T}\boldsymbol{Q}_i)\boldsymbol{1}
\end{aligned}$$

Similarly, the system of ordinary differential equations for disease propagation in outer-POD environment is

$$\boldsymbol{y}'_0 = (\boldsymbol{D} - \boldsymbol{I})\text{diag}(\boldsymbol{\mu})\boldsymbol{y}_0 + \text{diag}((\boldsymbol{D} + \boldsymbol{I})\text{diag}(\boldsymbol{\beta})\boldsymbol{y}_0)\boldsymbol{C}\boldsymbol{y}_0 - \frac{\text{diag}(\boldsymbol{\lambda})\boldsymbol{y}_0}{||\boldsymbol{y}_0||_1} + \text{diag}(\boldsymbol{T}\boldsymbol{Q}_0)\boldsymbol{1}$$

Combining the ODE systems for intra-POD and outer-POD disease propagation yields

the general modeling framework for the spread of infectious disease outbreak with medical operations and countermeasures. Notice that this modeling framework is derived based on the contact-based disease with only one POD. It can be easily generalized to model the scenario for vector-borne diseases and for multiple PODs in the region. For example, to model vector-borne disease, only the population dynamics due to disease propagation needs to be edited, since the modeling structure for medical operations and countermeasures remain the same regardless of the type of infectious disease. Therefore, the $\mathrm{d}_1\boldsymbol{y}_i$ term can be replaced by the differential equation for vector-borne diseases discussed in Chapter 1. To model the scenario with multiple heterogeneous PODs in the region, each of the POD can be modeled separately with the method discussed above and then aggregated according to the percentage of population served by that POD. Assume the region has M PODs, each serving p_i of the entire population, where $\sum_{i=1}^{M} p_i = 1$. Therefore, POD i in that region has a rate of arrival approximately equals to $p_i\mathrm{diag}(\boldsymbol{\lambda y}_0)/||\boldsymbol{y}_0||_1$. Additionally, since each POD in the region may have different layout design and throughput rate, different matrices $\boldsymbol{T}$ and $\boldsymbol{Q}$ that are specific to the settings of each POD can be assigned to the differential equations to model the heterogeneity.

2.4 Applications

In this section, we discuss two applications of the proposed modeling framework for human interventions and medical operations in containing the 2014 Ebola outbreak in West Africa and prioritizing vaccine administration for smallpox outbreaks.

2.4.1 Containing the 2014 West Africa Ebola Outbreak

The 2014 Ebola outbreak in West Africa is the largest ever of its kind. In this section, we derive a disease modeling and optimization framework to describe the development of the outbreak, estimate the future trend, and optimally allocate the medical resources available from the international aid to contain the outbreak. The study established a novel

compartmental model for Ebola virus and took the infections among healthcare workers into consideration. An optimization problem is formulated based on the results from the compartmental model to allocate the medical resources. The framework discussed in this study can also be applied to other epidemics to support the decision-making and emergency response procedure.

Introduction

Ebola virus is a zoonotic filovirus that causes a hemorrhagic fever syndrome in humans, with a high case-fatality rate. The 2014 West African Ebola virus outbreak is the largest ever to occur and has been proven difficult to control. The occurrence of cases in large urban centers enables the international spread via air travel, which may lead to global pandemics.

Mathematical models of infectious disease outbreak and epidemics are useful in utilizing the available information and transforming it into the knowledge of the disease while providing strategies and guidelines to prevent the disease from becoming a global pandemic. Disease modeling focuses on describing the mechanistic of the disease and models it as processes that result in the transition of individuals of the population between different stages. Fisman et al. applied the simple inclined decay with an exponential adjustment (IDEA) model to the epidemic and reproduced the observed pattern of disease growth. This model describes epidemic processes both in terms of exponential growth and simultaneous decay introduced by behavioral change, public health intervention, or any dynamic change that slows the disease transmission [88]. Compartmental model is among the most popular methods to describe the disease propagation. Althaus described this outbreak with an SEIR model and estimated the basic and the effective reproduction numbers in absence and presence of control interventions [10]. Legrand et al. proposed a stochastic compartmental model which classifies individuals into 6 stages and built the simulation based on this clas-

sification of individuals with different transition rate between stages [145]. Based on this model, Rivers et al. modeled five scenarios to examine their likely impact on the development of the epidemic, including improved contact tracing, decreased contact rate, increased proportion of hospitalized cases, and increased survival rate of hospitalized patients [185]. Chowell et al. described a transmission dynamics using an SEIR epidemiological model with quantified effects of social distancing interventions. The time-dependent transmission rate captures the effects of implementing public health interventions over time, which depicted the fact that the full effect of interventions is not seen immediately but gradually takes hold in the population [50]. Zaman et al. introduced optimal control strategies of a standard SIR model with time delay and use them to minimize the spread probability from the infectious population [217]. Drake et al. developed a discrete time, stochastic process model which comprises separate probability distributions for the number of secondary cases arising from health care workers (HCW) infected in hospitals, non-HCW infected by hospitalized patients, non-HCW infected from non-hospital care, and non-HCW infected through burial practices [70]. Besides the compartmental models, other methods to predict the outbreak trend include network models [127], stochastic simulation [205], and the Global Epidemic and Mobility model which captures the influence of potential traveling patients [98, 179]. These models were used to describe the disease propagation and estimate the potential impacts of Ebola virus.

Resource usage and allocation to contain the Ebola outbreak are also considered. Lewnard et al. developed a transmission model for Ebola virus and calibrated for the Ebola outbreak in Montserrado County, Liberia. Based on this model, they estimated the number of beds at healthcare facilities needed to effectively control the disease in Montserrado County is 4,800, which exceeds the 1,700 pledged by the USA to West Africa [147]. Bed occupation was also estimated by using the number of cases reported at each time point for the past ten Ebola virus outbreaks, while the maximum occupancy over the course of

the outbreak is obtained to give the bed per outbreak [32]. The impact of Ebola Treatment Centers (ETCs) capacity (such as the insufficient number of beds) and admission rate on the hospital-based interventions and the early and drastic enhancement of infection control measures in healthcare settings are also analyzed [171, 174].

Besides the characteristics of the Ebola virus and lack of public health infrastructure, socio-culture in West Africa where the outbreaks take place also plays an important role and complicated the implementation of control interventions. Funeral attendants could facilitate the development of major epidemics since the local cultural practices involve touching the body of the deceased [17]. Lack of knowledge of the disease leads communities to deny its existence and associate it with witchcraft or conspiracy theories [47]. These create barriers for international intervention aiming to slow down the spread of Ebola.

In this study, we propose a compartmental model-based mathematical model to describe the spread of Ebola virus in multiple West African countries during the 2014 outbreak in both civilian and healthcare worker population. Additional stages including hospitalization and funeral are included to characterize the special properties of Ebola virus. Then we formulate an optimization problem on top of the output of the ordinary differential equations and solve it with a line search algorithm to find the minimum number of beds in healthcare facilities needed to achieve early containment in each country. We test our model using the real data from Guinea, Sierra Leone, and Liberia and compare the projected trend of the outbreak under different scenarios with the actual reported cases. Our result suggests that optimal allocation of limited medical resources during pandemics is crucial to minimize the cumulative infections and achieve early containment.

Methods and Designs

Based on the property of the Ebola virus and the availability of medical resources in West Africa back in 2014, we propose an ordinary differential equation-based disease propagation model that generalizes the SEIR model for Ebola. Additional stages are added to depict the disease development of Ebola virus. Due to the high possibility of transmission during the traditional funeral without protection, dead patients that do not receive cremation will still be infectious. A new stage F for funerals is added to model the transmission happens during funerals. If a dead patient is buried instead of receiving cremation, it is not removed from the system immediately; instead, it stays in the stage F for a certain period before being removed. During that period, it is able to infect others, and this period is the funeral duration. Another new stage introduced to our model is H, which stands for hospitalization. Infected individuals can get hospitalized with some probability (depending on the availability of the medical resource), and they are quarantined from other civilians. Therefore, patients in stage H are not infectious to civilians but are infectious to healthcare workers that have direct contact with them. Similar to stage I, a patient in stage H can either recover (R) or die, and the dead hospitalized patients can either be buried or cremated. Therefore, the system for civilians contains seven stages: S (susceptible), E (exposed), I (infectious), R (recovered), H (hospitalized), F (funeral), and D (deceased; removed from system). Let $N = S + E + I + R + F$ denote the total civilian population in the system.

The seven-stage model described above models the virus transmission among civilians. The healthcare workers in West Africa are also in danger of getting infected, and they face the different situations as the local civilians: they have proper protection, thus the possibility of getting infected is lower than the civilians; they have enough resource such that all infectious healthcare workers can get treated within a healthcare facility and get quarantined from other people; the dead healthcare workers are dealt with caution so that they do not infect others, i.e., there is no F stage for the healthcare workers. Let

S_H, E_H, I_H, H_H, R_H, and D_H denote the six corresponding stages for healthcare workers. Let $N_H = S_H + E_H + I_H + R_H$ denote the total population of healthcare workers in the system. Notice that the two groups above are not isolated. Healthcare workers get exposed (E_H) because of their physical contact with the patients being hospitalized (H), as well as the infectious healthcare workers (I_H). In this study, the healthcare workers are assumed to stay inside the healthcare facilities all the time so that they do not get infected through contacting infectious civilians (I), and the infected healthcare workers (I_H) do not infect healthy civilians (S). However, these assumptions can be relaxed by using the time-variant parameters discussed later.

Let μ_S denote the average number of infectious contacts each individual makes in a unit time, i.e., the baseline contact or infection rate; μ_E is the transition rate from exposed stage to the infectious stage and it is characterized by the incubation period of Ebola virus; μ_I is the transition rate from infectious stage to stage R, F, and D and it is determined by the duration of the infection; μ_{IH} is the transition rate from infectious stage to hospitalized or the inverse of the time between infection to admission to healthcare facility; μ_H is the rate from the hospitalization stage to other stages or the inverse of the duration of hospitalization; μ_F is the transition rate from stage F to stage D and it is characterized by the duration of the funeral.

Other parameters in the model include H_m, the total number of beds available. This reflects the availability of healthcare resource. p_{HD} and p_D are the death rate with and without hospitalization. p_{HF} and p_F are the percentage of funerals (non-cremation) for deceased patients with or without hospitalization, respectively. p_{FS} is the adjusted factor of funeral infection rate, and f is the factor of healthcare worker transmission rate, reflecting the protection of healthcare workers. Plugging these parameters to the modeling framework

discussed in Equation 2.1 yields the following system of ordinary differential equations:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}S &= -\mu_S(I + p_{FS}F)S & \frac{\mathrm{d}}{\mathrm{d}t}S_H &= -f\mu_S(H + I_H + H_H)S_H \\
\frac{\mathrm{d}}{\mathrm{d}t}E &= \mu_S(I + p_{FS}F)S - \mu_E E & \frac{\mathrm{d}}{\mathrm{d}t}E_H &= f\mu_S(H + I_H + H_H)S_H - \mu_E E_H \\
\frac{\mathrm{d}}{\mathrm{d}t}I &= \mu_E E - p_H\mu_{IH}I - \mu_I I & \frac{\mathrm{d}}{\mathrm{d}t}I_H &= \mu_E E_H - \mu_{IH}I_H \\
\frac{\mathrm{d}}{\mathrm{d}t}H &= p_H\mu_{IH}I - \mu_H H & \frac{\mathrm{d}}{\mathrm{d}t}H_H &= \mu_{IH}I_H - \mu_H H_H \\
\frac{\mathrm{d}}{\mathrm{d}t}R &= (1 - p_D)\mu_I I + (1 - p_{HD})\mu_H H & \frac{\mathrm{d}}{\mathrm{d}t}R_H &= (1 - p_{HD})\mu_H H_H \\
\frac{\mathrm{d}}{\mathrm{d}t}F &= p_F p_D \mu_I I + p_{HF}p_{HD}\mu_H H - \mu_F F & \frac{\mathrm{d}}{\mathrm{d}t}D_H &= p_{HD}\mu_H H_H \\
\frac{\mathrm{d}}{\mathrm{d}t}D &= (1 - p_F)p_D\mu_I I & & \\
&\quad + (1 - p_{HF})p_{HD}\mu_H H + \mu_F F & &
\end{aligned}$$

The model above divides the spread of disease into two types: the disease propagation that happens outside the healthcare facilities and infects civilians, and the disease propagation within the healthcare facilities that infects healthcare workers. In addition, heterogeneous infectivity among civilians and healthcare workers is characterized by the factor f to capture the proper protections of the healthcare workers. In addition, this model is not limited to the Ebola outbreak in West Africa; by setting proper parameters, it can be used to estimate the virus spread in any other country or predict the potential trend of outbreaks of other types of contact-based diseases.

One major challenge faced by West African countries is the lack of healthcare facilities and infrastructures. In particular, during the Ebola outbreak, the West African countries like Guinea does not have enough beds in their hospitals or healthcare facilities, thus not all patients can be quarantined or treated by healthcare workers. When there are no available beds, patients can only be treated at home, which brings risks to their relatives and neighbors since their infectivity is much higher than the healthcare workers. The parame-

ter p_H in the model above can be used to capture the available beds in a country: if there are beds available at hospitals or healthcare facilities, then $p_H = 1$, i.e., infectious patients will be admitted to quarantine; otherwise $p_H = 0$. Essentially, p_H can be seen as an indicator of whether there are vacancies for quarantine. Therefore, $p_H = \mathcal{I}\{H < H_m\}$, where $\mathcal{I}$ is the indicator function. We use the number of beds available in each country as the main factor in controlling the spread of the virus and investigate the impact of availability and delayed supply of beds to determine the number of beds required to contain the outbreak in each country. Then we estimate the total infections through the end of 2015, including those of healthcare workers. These additional beds and the corresponding resources and healthcare workers are from international assistance. However, the international assistance is limited, and devoting too many resources to the infected area does not necessarily accelerates the containment, but increases the risk of infections among healthcare workers. Therefore, we propose an optimization framework to minimize the number of beds required, or in general, the medical resources, subject to the constraints on the total infections among civilians and healthcare workers and the containment time. In particular,

$$\min H_m \quad \text{s.t.} \quad I^* \leq I_0, I_H^* \leq I_{H0}, T^* \leq T_0$$

where I^* and I_H^* are the numbers of infections among civilians and healthcare workers by the end of 2015, respectively, while I_0 and I_{H0} being the target number of infections among these two population groups at the end of 2015. T_0 is is the target time of containment, and T^* is the actual time of containment, defined as:

$$T^* = \min \left\{ t : \frac{I_{t+\epsilon} - I_t}{I_t} \leq \delta \right\}$$

i.e., the earliest time such that the increment rate of infections is controlled under a certain level δ. Notice that other constraints can also be introduced to obtain the optimal allocation of healthcare resources.

The variables in this optimization framework are the output of the system of ordinary differential equations shown earlier, thus searching for optimal solutions is computationally expensive. We apply a line search algorithm for a given structure of parameters and search for the optimal setting iteratively on the surface generated by the output of the differential equations. Each iteration requires about 10 CPU minutes to solve.

Results

Base on the model proposed in the previous section, numerical experiments are performed on three influenced countries: Guinea, Sierra Leone, and Liberia to validate the model and predict the trend of the Ebola outbreak. The parameters that are common in all experiments are listed in Table 2.1.

Table 2.1: Ebola: model parameters used in all numerical experiments

Parameter	Value	Literature Source
Incubation period ($1/\mu_E$)	Uniform distribution between 8 and 12 days	[76, 209]
Infection duration ($1/\mu_I$)	Uniform distribution between 5 and 9 days	[209, 210]
Hospitalization duration ($1/\mu_H$)	Uniform distribution between 5 and 15 days	[209]
Funeral duration ($1/\mu_F$)	1 day	[170]
Percentage of cremation ($1 - p_F$)	40% (70% after Dec 1, 2014)	[170]
Percentage of cremation if hospitalized ($1 - p_{HF}$)	80%	[170]
Fatality rate (p_D)	70%	[209, 210]
Fatality rate if hospitalized (p_{HD})	50%	[209, 210]
Adjustment factor of healthcare worker transmission rate (f)	0.05	[213]

Other parameters vary between countries, and the basic reproduction number for each country is estimated using the historical statistics reported in each country. These parameters are listed in Table 2.2.

Table 2.2: Ebola: model parameters specific to each country

Parameter	Value			Literature Source
	Guinea	Sierra Leone	Liberia	
Basic reproduction number ($\mathcal{R}_0$)	1.45	1.4	1.45	Fitted
Mean time from infection to hospitalization ($1/\mu_{IH}$)	$3-4$ days	4 days	4 days	[209]
Number of healthcare workers as of Oct 1, 2014	300	$1,000$	500	[212]

The number of projected infections in Guinea under different scenarios is fitted and the result is shown in Figure 2.3. The intervention effect was observed starting from April 15. Changing the basic reproduction number $\mathcal{R}_0$ twice (1.45 to 0.9, then to 1.3) during the intervention reflects the change of human behaviors and the effort of interventions with public awareness of the disease outbreak. If the unreported cases are taken into consideration, this model will have more predicting power. If a total of 50 beds were available in middle April, the epidemic could have been rapidly contained in middle 2014 with no more than 250 total infections. According to the WHO report, Guinea had 160 beds in October 2014. If this situation did not change since then, the epidemic would be contained in the middle of 2015 and would result in about 3,800 infections by the end of 2015. If 200 beds were available by November 2014, the outbreak would contain by the end of 2015 and the cumulative infections would be about 3,500.

The number of infections in Sierra Leone under different scenarios is fitted and the result is shown in Figure 2.4. The intervention started on August 1, 2014. According to the WHO report, 130 beds and 300 beds were available on September 5, 2014, and September 17, 2014, respectively. If no additional beds were available since then, the epidemic would contain in middle 2015 and the total infections by then end of 2015 would be 14,200. If 350 beds or 500 beds were available at the beginning of 2015, the epidemic would contain in the early 2015 and the total predicted infections would be 12,100 and 11,000, respectively.

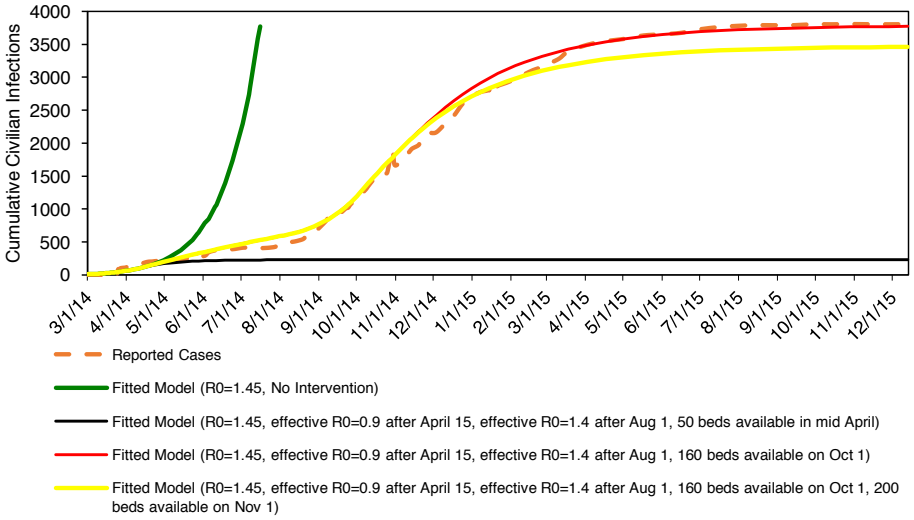


Figure 2.3: Ebola: infections in Guinea under different scenarios.

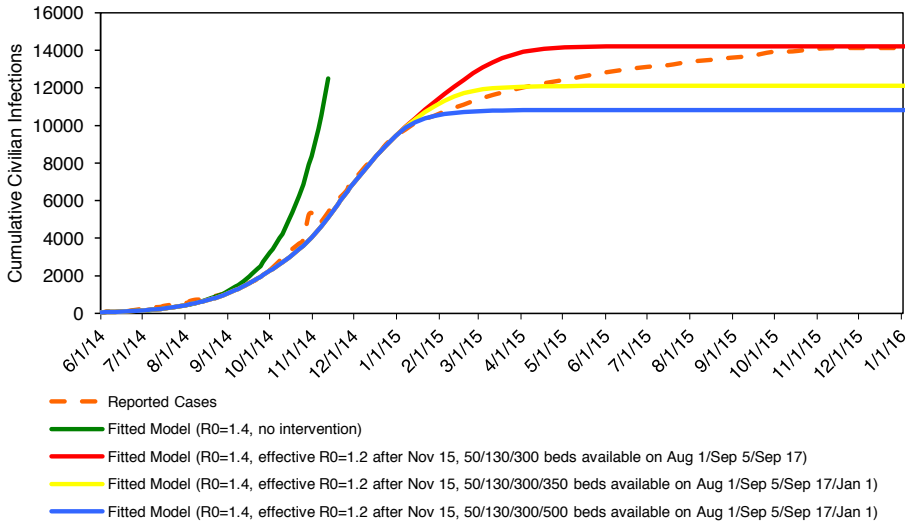


Figure 2.4: Ebola: infections in Sierra Leone under different scenarios.

The number of infections in Liberia under different scenarios is fitted and the result is shown in Figure 2.5. The intervention started in September 2014. If the number of beds available at the beginning of September 2014 was 500, then the epidemic would contain in the early 2015 and the total infections would be 6,000. The number of beds available in the early October 2014 was essential to the development of the disease. If 600 beds were available on October 3, 2014, the predicted number of infection at the end of 2015 would be 10,300; if these beds were delayed for two days, this number would become 12,700;

if 20 more beds were available, the cumulative infections at containment would be 9,500. Therefore, the number of beds available at the fast-spreading phase of the outbreak has a huge impact on its development.

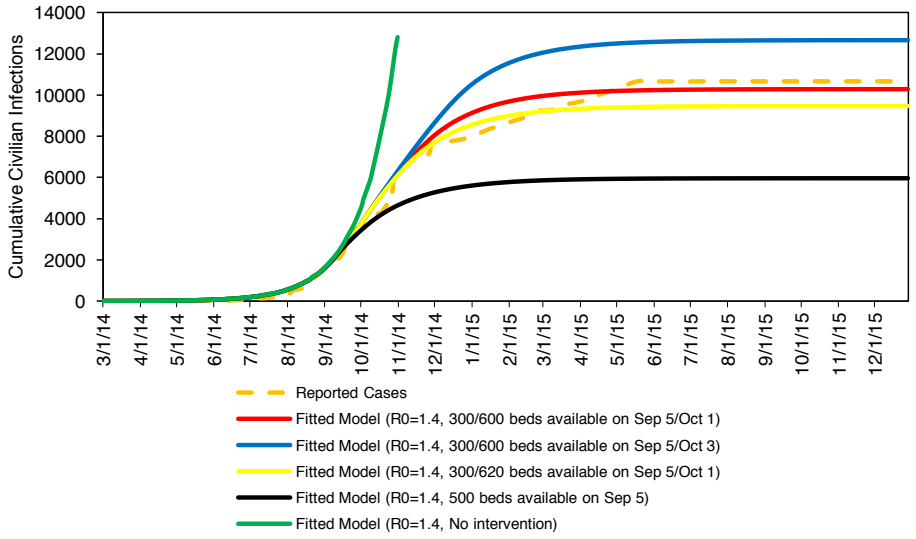


Figure 2.5: Ebola: infections in Liberia under different scenarios.

The number of infections among healthcare workers in the three countries is shown in Figure 2.6. Though they have proper protection against the virus, the healthcare workers are still in danger of infection due to their high contact rate to infectious populations. At the initial phase of the pandemic, i.e., the second half of 2014, the number of infections among healthcare workers increases drastically. For example, the infections among healthcare workers in Liberia reached 250 after 6 months of the initial outbreak according to the prediction. As the epidemic contains, the number of infections among healthcare workers also stabilizes. There are several reasons to explain the rapid increment of infections among healthcare workers: insufficiency of personal protective equipment before the international interference, the improper use of the protective equipment, and too few medical staff for the large outbreak at this scale. We should notice that the loss of medical staff including experienced doctors and nurses has made it increasingly difficult to control the outbreak, and the international community has to ensure a sufficient number of foreign healthcare

workers. This further impedes the containment of this outbreak.

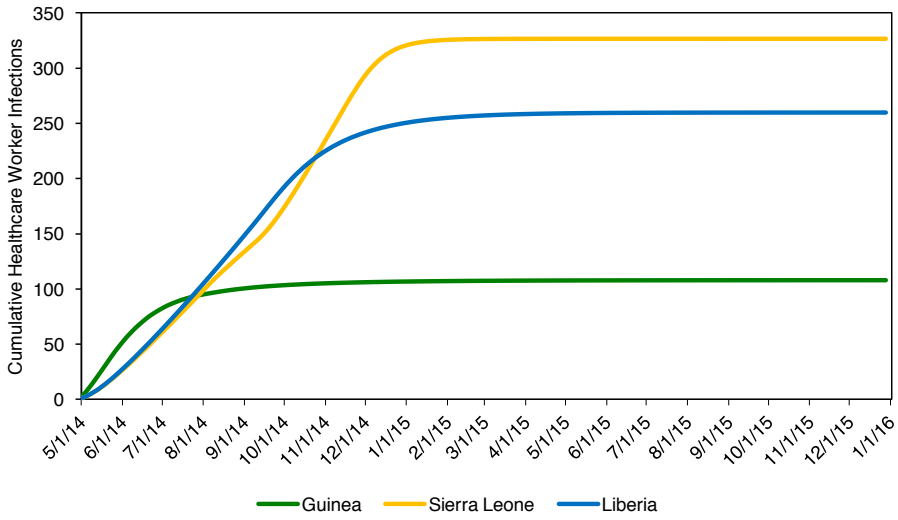


Figure 2.6: Ebola: infections among healthcare workers in 3 countries.

Discussion

In this study, we proposed a compartmental model to describe the spread of Ebola virus in 3 West African countries during the 2014 outbreak. This model consists of two parts: the compartment model for civilians in West Africa, and the compartment model for healthcare workers. We fitted the model based on the number of infections in Guinea, Sierra Leone, and Liberia reported by CDC and predicted the trend of Ebola outbreak by the end of 2015. It is estimated that the outbreak will contain by the end of 2015 given enough beds in these countries, i.e., enough availability in hospitals and healthcare facilities so that most infected civilians can get quarantined. If there were enough beds since the initial phase, the outbreak could have been contained early; also, if the transition rate from stage I to stage H is faster, which requires the rapid diagnosis of infectious patients and immediate transition to hospitalization, the outbreak would also contain faster. The infections among healthcare workers in the affected countries are also estimated. Though equipped with the proper protections, the healthcare workers still are at risk of getting infected and this is

extremely dangerous if they return to their countries.

The proposed model is general-purpose and can be easily extended to accommodate more complicated scenarios, such as multiple groups with different parameters. In this model, the baseline contact rate between civilians is treated as a constant. Since the Ebola virus spreads via direct contact with the blood, secretions, organs or bodily fluids, the possibility of infection rate of each individual in stage S may be different regarding how they interact with the infectious individuals, or whether they make direct contact. Under this setup, the civilians can be further divided into two groups and different parameters can be assigned to differentiate these two different contacts. Similarly, as an increasing number of organizations are involving in this outbreak, other groups such as volunteers and military forces can be added to the model with their own compartments and model parameters. This enables us to predict the number of infections among different groups and evaluates their risk in West Africa respectively.

Another advantage of the proposed model is the capability of capturing human behavior through the horizon of the outbreak. Instead of assigning each parameter a constant value, we may set the parameters as functions of time. During different times of a day, people act differently and this affects the spread of the virus. For example, the healthcare workers may be out of the healthcare facilities or hospitals at the certain time of a day, and this makes it possible that a civilian not in quarantine will make direct contact with a healthcare worker. If the healthcare worker is infected but not hospitalized yet, the civilians may also get infected. For this scenario, we can modify the ordinary differential equation systems with the time-variant parameters. For example,

$$\frac{\mathrm{d}}{\mathrm{d}t}S = -\mu_S \frac{N}{N + \eta(t) p_O N_H}(I + p_{FS} F \eta(t) p_O I_H) S$$

Where $\eta(t)$ is a time-variant parameter indicating if the healthcare workers are out of the healthcare facilities at time t, and p_O is the proportion of the healthcare workers that are out. Similarly, the healthcare workers that are outside may have direct contacts with the infected civilians or the unsafe funeral. Then

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}S_H = &-(1-\eta(t))f\mu_S(H+I_H+H_H)S_H \\ &-\eta(t)f\mu_S\left[\frac{N_H+H_H}{(1-p_O)N_H+H_H}(H+(1-p_O)I_H+H_H)(1-p_O)S_H\right. \\ &\left.+\frac{N_H+H_H}{p_ON_H+N}(I+p_{FS}F+p_OI_H)p_OS_H\right]\end{aligned}$$

This equation captures the behavior of healthcare workers during different times of the day. Other equations in the system of ordinary differential equation can be updated accordingly. Other parameters such as the basic reproduction rate can also be defined as functions of time to accommodate the changes in policy or environment during the outbreak. This enables the modeling and analysis of more complicated situations.

The disease modeling and optimization framework proposed in this study can be adapted to other kinds of epidemics. With proper set up of the compartmental model and parameters, as well as the definition of medical resources and the role they play in controlling the spread of the disease, one can apply the similar methodology and analysis procedure to predict the future development of the outbreak and estimate the optimal allocation of medical resources to contain the outbreak. This will be of great significance in the future preparedness of global pandemic.

2.4.2 Prioritized Vaccination Strategy to Contain Smallpox Outbreaks

When the vaccine supply is limited during infectious disease outbreaks, prioritized vaccination is broadly considered as an effective strategy. In this section, we propose a

generic framework to model the effect of mass vaccination during pandemics. This modeling framework integrates the compartmental model for disease propagation and stochastic queuing model for vaccination operations. We applied the model to a smallpox outbreak and derived the optimal trigger for switching from prioritized to non-prioritized strategy numerically to minimize the overall attack rate and mortality rate under different scenarios. The value of optimal switching trigger is closely related to the vaccine supply level, the composition of the population, and the properties of the disease. The effect of optimal mixed strategy is also governed by multiple model parameters including the delay of vaccine supply, dispensing efficiency, and accuracy of triage. Therefore, to maximize the protective effect of available vaccines, determining the optimal mixed strategy and implementing the strategy effectively during on the ground operations are equally important. This model can be easily generalized to solve the optimal switch trigger between strategies for other types of disease outbreaks by supplying the corresponding disease propagation structure and parameters. The system enables decision makers to find the best vaccination strategy in real time to optimally utilize the limited medical resource and achieve faster containments during pandemics.

Introduction

Vaccination is an important method for controlling the spread of infectious diseases worldwide. However, the development and manufacturing of vaccines usually take months to years, and vaccine shortages during epidemic outbreaks are not uncommon [168, 201]. The recent yellow fever outbreaks in Brazil and multiple African countries also exposed the insufficiency in vaccine supply in case of large-scale infectious disease outbreaks [95, 129]. The condition of the outbreak is difficult to control without a continuous supply of vaccines. On the other hand, Preparing and stocking vaccines for the population in the entire U.S. is financially impractical [189]. The incapability of supply chain system in some developing countries impedes the distribution of vaccines as well [120]. Because of

the limited supply of vaccines, the effect of reduced and diluted dosage has been studied [91, 132], aiming at covering more population with the current supply level of vaccines. In addition, individual-level vaccine immunogenicity prediction using system biology and machine learning is recently studied to improve the effect of vaccination [140, 181].

Besides the effort to best utilize vaccines at individual-level, systematically prioritized vaccination is widely considered one of the best strategy to contain a pandemic [161, 202]. Normally, healthcare workers, children, pregnant women, and individuals with special medical conditions will be classified as the high-risk group and be the first batch to receive the vaccines [178]. The proper definition of the high-risk group who receive vaccines first is also extensively discussed [56, 77]. Since searching for the optimal prioritized strategy for a general definition of risk groups under practical constraints can be computationally difficult, some studies solved a relaxation of the problem by assuming adequate vaccine supply [199] or prioritizing a relatively simply-structured risk group [54]. Other works on prioritized vaccination strategy include analysis fusing the spatial information with demographic data [215], minimizing the year of life lost [162], and optimizing multiple evaluation metrics [160].

Wallinga et al. studied the optimal allocation of medical resources with limited data by targeting intervention measures at the group with the highest risk of infection per individual [206]. Mylius et al. studied the vaccination strategy under different supply scenarios of influenza vaccines and suggested prioritizing individuals with high-risk of complications [166]. Patel et al. used stochastic epidemic simulations and meta-heuristic to find the optimal vaccine distributions [176]. Computer simulation is also used to compare different prioritization strategies and definitions of risk groups to assist decision-making, as discussed in Lee et al. [133, 134]. While most studies assume the supply of vaccines is instant, Meyers et al. optimized the allocation of vaccines when there is a delay in vaccine

supply [159]. In practice, most public health departments in the U.S. first use a prioritized strategy to cover high-risk population and switch to a non-prioritized strategy later. Lee et al. proposed a compartmental-based model to determine the optimal switch trigger between prioritized and non-prioritized strategies [143] and solved the problem of determining the switch point via numerical optimization.

In this study, we focus on the prioritized vaccination strategy to mitigate smallpox outbreaks. Although smallpox has been eliminated since the 1970s and vaccinations are no longer widely dispensed since then [85], it remains a threat to public health due to the possibility of bioterrorism, and the preparation of mass vaccination events is necessary [6, 183]. Clinical researches suggest that smallpox vaccines may cause severe adverse reactions [42]. However, dispensing smallpox vaccines, either targeted or mass vaccination, is still considered as the most effective method of containing the outbreak [106]. Multiple studies have already been carried out to identify the proper vaccination strategies in case of a smallpox outbreak to achieve early containment and minimize the total infections [31, 87, 118].

We consider the mass vaccination strategy after the initial outbreak of smallpox using the same methodology described in [143] to find the optimal switch trigger between prioritized and non-prioritized strategies. We develop a generic modeling framework for human intervention, especially the effect of mass vaccination, on the disease propagation process in compartmental models. The vaccination operations inside the vaccine clinics or point-of-dispensing sites (POD) are explicitly modeled. Then we use numerical optimizations to find the optimal switch trigger in terms of minimizing overall attack and mortality rate under different levels of vaccine supplies and severity of the outbreak (determined by the basic reproduction number and the initial infectious population). Agent-based simulation is used to validate our result. Sensitivity analyses are then performed on the proposed

modeling framework and the robustness of the conclusion is examined. This modeling framework not only applies to mitigating smallpox outbreaks but can assist the decision-making process in various infectious disease outbreaks as well with correct compartmental model plugged in. These properties make the model has practical attractions. Healthcare departments and decision makers can input multiple model parameters with respect to the biological feature of the disease, vaccine supply, demographic information, etc., and use the suggested optimal switch trigger to facilitate the vaccine dispensing in emergencies to minimize the infections and mortalities due to infections.

Methods and Designs

To study the transmission of smallpox in the population, we develop the baseline compartmental model with 6 compartments: susceptible (S), exposed (E), incubation period (P), infectious (I), recovered and immune (R), and deceased (D). Besides the change of compartments due to disease propagation and new infections, the vaccination process is also modeled in detail to reflect the importance of vaccination in controlling smallpox outbreaks. We use the intra-POD and outer-POD disease propagations discussed in Section 2.3 to model the spread of the disease inside and outside PODs. The existence of POD and vaccines adds two more compartments to our model: hospitalization (H) for visiting individuals with symptoms of smallpox infections and will be properly treated and quarantined, and vaccinated and immune (V) for visiting individuals who received vaccines and developed immunity. Figure 2.7 shows the stage transition diagram of this model, where the solid lines are transitions associated with new infections and disease propagation and the dashed lines are transitions associated with the vaccination and treatment.

Following up the model in Section 2.3, multiple PODs will be set up for dispensing vaccines, and quarantine and treat infectious individuals in case of disease outbreaks. We use the POD setup following the flowchart shown in Figure 2.8 to model the operations

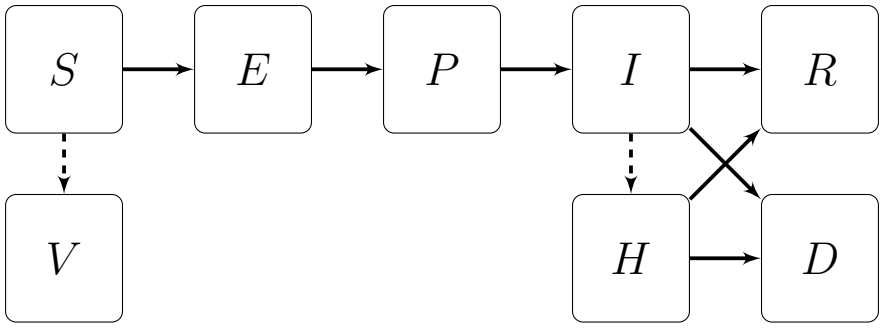


Figure 2.7: Smallpox: transition diagram with vaccination and hospitalization.

of PODs. Individuals at all stages may arrive at the POD, while only the asymptomatic ones (susceptible, exposed, and incubation period) will receive vaccines, and we assume that only the susceptible (S) and exposed (E) individuals may develop immunity after vaccination based on previous researches on smallpox virus with a probability [109, 155].

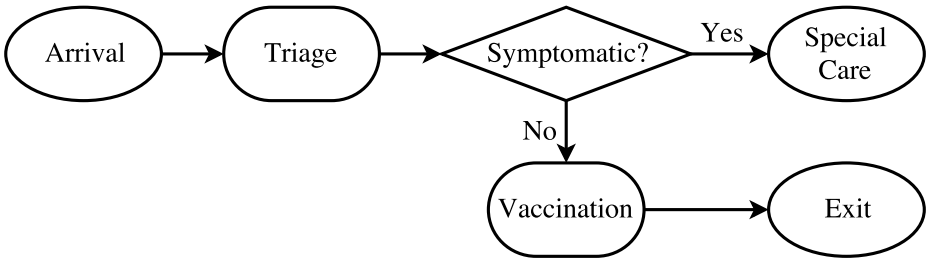


Figure 2.8: Smallpox: flowchart for dispensing vaccines during an outbreak.

To demonstrate the importance of prioritizing vaccination, we separate the population into five different risk groups: normal adults, children under 10 years old, pregnant women, patients with associated diseases, and healthcare workers. The latter four groups are categorized as high-risk groups: healthcare workers have higher contact rates to infections; children, pregnant women and patients with associated health conditions are more vulnerable to infections [192]. In addition, since routine vaccination in the U.S. against smallpox stopped in 1972 and worldwide in 1984 [85, 169], the younger individuals in the population can also be prioritized should it be permitted. In the rest of this paper, we will use the risk group model with five groups to discuss the prioritization of vaccines.

To better reflect the realistic and typical emergency response in case of a pandemic outbreak, we assume that the arrival rate of the PODs is a time-variant parameter instead of a constant throughout the process of the disease outbreak. Let λ denote the baseline arrival rate of individuals, i.e., the arrival rate at the beginning of the vaccination process if the non-prioritized policy is used. Then if prioritized policy is used, all the arrivals are from the high-risk group, with population-adjusted arrival rate:

$$\lambda(t) = \lambda \frac{N_0^H(t)}{N^H}$$

where N^H is the total population of the high-risk group, and $N_0^H(t)$ is the amount of the high-risk population at time t who have not visited the POD yet. With the progress of vaccine administration, the arrival rate at PODs will gradually decrease as more individuals are vaccinated or treated. If we switch to the non-prioritized policy, the expression of arrival rate will become

$$\lambda(t) = \lambda \frac{N_0(t)}{N_0}$$

with N and $N_0(t)$ being the initial population and population not visited POD in both the high-risk and normal group. The basic arrival rate λ is inferred from our experience in past time-motion studies, and the composition of population is obtained using the demographic statistics. Based on the expression of instant arrival rate, there will be an increase in the arrival rate upon the switching, as the high-risk population is significantly smaller compared to the normal group. The switch happens if the percentage of vaccines administrated to the high-risk group reaches the switch trigger g. In another word, let π be the percentage of high-risk population, and let c denote the percentage of coverage with the initial vaccine inventory. Then the switch happens if gc/π of the high-risk population are protected. Therefore, if $g = 0$, then the corresponding strategy is full-non-prioritized, i.e., the vaccines are administrated to everyone regardless of their risk group; on the other hand, if $c \leq g \leq 1$, then the associated strategy is full-prioritized, where all vaccines will be

distributed to high-risk groups, and the normal group will not receive vaccines.

We use AR, the overall attack rate, as our evaluation metric. Overall attack rate measures the percentage of the population who experience the smallpox infection during the horizon of the outbreak. It is the cumulative number of infections throughout the outbreak. Then we seek the optimal switch trigger g^* for the mixed strategy that minimizes the overall attack rate by the end of the outbreak with constraints on the availability of vaccines. In our example, we assume that there is no replenishment of vaccine inventory. This can be relaxed by introducing additional parameters to capture the arrival batches of vaccines [143]. The objective function AR is derived from the output of the ODE system which is controlled by the switch trigger g. Therefore, to obtain the optimal value g^*, we construct multiple ODE systems by varying the value of g and run the system until equilibrium is reached, then apply a line search algorithm to find the optimal solution that minimizes AR. Due to the complexity and the nonlinearity of the ODE systems, each instance of switch trigger optimization requires about 6,000 CPU minutes to solve.

The analyses we performed are based on the demographic statistics of the state of Georgia with 10.21 million population, among which 19.30% are high-risk population (14.8% children under 10 years old, 3.0% healthcare workers, 1.3% pregnant women, and 0.2% patients with associated disease conditions). 43 PODs will need to be set up across the state to provide vaccination and medical assistance for the entire region. Due to the limited supply of smallpox vaccines, we assume that each individual will only receive one dose of vaccine. These data will be supplied to the modeling framework proposed in Section 2.3 to establish the system of ordinary differential equations to describe the dynamics of the system during the smallpox outbreak. The parameters used in the simulation are summarized in Table 2.3.

Table 2.3: Smallpox: model parameters used for analysis.

Parameter	Value	Literature Source
Distribution of population in different risk groups	$[80.7\%, 14.8\%, 3.0\%, 1.3\%, 0.2\%]$	World Bank
Basic reproduction number of smallpox	5 to 7	[75]
Percentage of initial infectious population	0.5% to 2.0%	Assumed
Mean dwelling time of exposed stage	5 days	[85, 109]
Mean dwelling time of incubation period	15 days	[85]
Mean dwelling time of infectious stage	20 days	[85]
Mean time between showing symptoms to hospitalization/quarantine	1 day	Assumed
Mean hospitalization time	10 days	Assumed
Probability of vaccine effectiveness	90%	[15]
Mortality rate without treatment	5%	[85]
Risk factor of infection in different groups	$[1, 2, 1, 3, 3]$	Assumed
Initial baseline arrival rate	15/min	Assumed

Our modeling framework is highly flexible and can be extended in multiple ways. Thus, any configuration of the POD layout and any structure of disease propagation can be incorporated. Also, with the computational capability of RealOpt-POD [141], the large-scale modeling and optimization system for public health infrastructure developed at CDC, we use agent-based simulation to trace the behavior and status of every client inside the PODs and use this result can to compare and validate the output of the ODE systems.

Results

We first demonstrate the result of optimal switch triggers for the mixed vaccination strategy under three different scenarios: 20% vaccine supply, 30% vaccine supply, and 40% vaccine supply, where the supply level indicates the percentage of total population that is covered. Figure 2.9 shows the result of the overall attack rate and mortality rate against different high-risk group coverage under the typical basic reproduction number of smallpox virus $\mathcal{R}_0 = 5$ and initial infectious population set to 1%. The overall attack rate and mortality rate are evaluated on day 270 after the initial outbreak. The optimal switch triggers and the associated high-risk group population coverage under the three different supply

scenarios are marked. Notice that based on our assumption of the total population, a 1% reduction in attack rate results directly to over 100,000 reductions in infectious individuals, demonstrating the importance of applying the mixed strategy with optimal switch trigger. In addition, the curvatures of all three scenarios tend to reach their maxima around the optimal trigger. This indicates that the selection of optimal switch trigger is rather robust and can easily be translated into practical implementations with large fault tolerance. In case of a real emergency, the inventory data of vaccines, the calculation of proportions of high-risk populations, or the record of the administrated vaccines may not be accurate. Then the robustness of this model is critical and desirable.

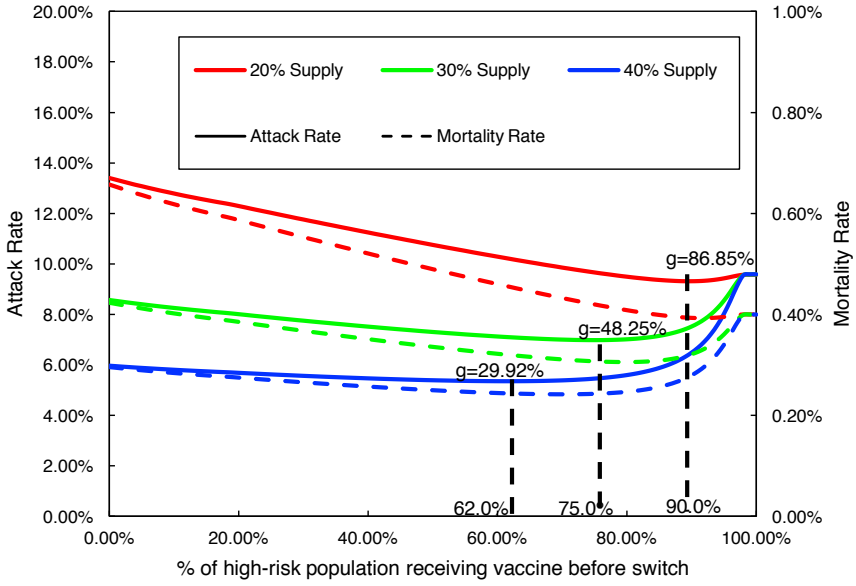


Figure 2.9: Smallpox: attack and mortality rate under different vaccination strategies.

Under the typical basic reproduction number of smallpox virus $\mathcal{R}_0 = 5$, the optimal mixed strategy is superior to either non-prioritized or non-full-prioritized strategy. The attack rates of three scenarios with the full-non-prioritized strategy are 13.4%, 8.56%, and 5.97%, respectively; the attack rates with full-prioritized strategy all converge to 9.58%, as the vaccine supply can cover the entire high-risk population in all three scenarios. The optimal mixed strategy, on the other hand, generates 9.32%, 6.97%, 5.35% attack rates in

three scenarios.

Table 2.4: Smallpox: attack rates of vaccination strategies under different scenarios

$\mathcal{R}_0$	Vaccine supply	Optimal mixed strategy		% increase in attack rate	
		Switch trigger	Attack rate	Full-prioritized	Full-non-prioritized
5	20%	86.85%	9.32%	+2.79%	+43.78%
	30%	48.25%	6.97%	+37.56%	+22.81%
	40%	29.92%	5.35%	+79.07%	+11.59%
6	20%	85.88%	19.02%	+3.71%	+43.51%
	30%	47.54%	12.74%	+55.18%	+32.18%
	40%	29.14%	8.21%	+135.55%	+18.87%
7	20%	84.81%	32.96%	+4.11%	+24.90%
	30%	46.27%	22.37%	+53.40%	+26.34%
	40%	28.98%	13.80%	+145.84%	+22.27%

Table 2.4 contrasts the optimal switch trigger for mixed strategy under different combinations of vaccine supply level and basic reproduction number $\mathcal{R}_0$. When $\mathcal{R}_0$ increases, the optimal switch trigger for all supply levels drops, indicating that the time for implementing non-prioritized strategy is earlier. This observation is intuitive since the more severe the pandemic is, the earlier that the strategy protects all population instead of only the high-risk population, as the high basic reproduction number tends to fade out the effect of high infectivity associated with high-risk population. Results also suggest that the overall attack rate evaluated at the containment of the optimal mixed strategy is much lower compared to other two strategies and the optimal mixed strategy can halve the infections throughout the outbreak in some cases (see 40% supply when $\mathcal{R}_0 = 6$). In addition, we observed that by using the optimal mixed strategy, the outcome of cumulative attack rate can be superior to that of non-optimal strategies with higher vaccine supply level. This as well addresses the significance of optimal control in vaccine distribution.

Figure 2.10 shows the attack rate against the number of days after the initial outbreak under three different strategies: optimal mixed strategy, full-prioritized strategy, and full-

non-prioritized strategy when the vaccine supply level is 40% and $\mathcal{R}_0 = 5$. Individual-based stochastic simulation (displayed in solid lines) is used to verify the output of the ODE systems (displayed in dashed lines). Both methods suggest the same result: optimal mixed strategy significantly reduces the attack rate throughout the outbreak and leads to an early containment. For example, with full-prioritized strategy, the attack rate at containment is above 10% and the outbreak starts to contain at around Day 200. The optimal mixed strategy can achieve an overall attack rate of 5.35% (465,000 in population) and bring the time of containment to Day 150, a 25% improvement compared to the full-prioritized strategy.

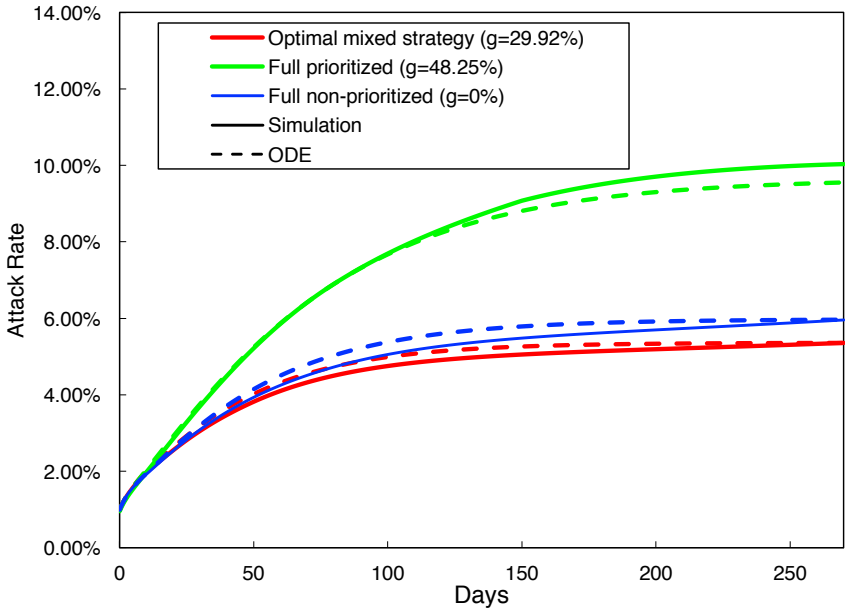


Figure 2.10: Smallpox: attack rate under 3 vaccination strategies with ODE and simulation.

Figure 2.11 compares the effect of delay in supply of vaccines, when $\mathcal{R}_0 = 5$, initial infection is 1%, and the vaccine supply level is 40% dispensed with the optimal mixed strategy. In case of no delay, the peak of the daily prevalence of disease appears at Day 15 with 1.60% maximum and starts to reduce afterward. If the vaccine supply is delayed by 3 weeks, the number of infectious population will reach a peak at Day 30 with 2.68% max-

imal prevalence. The longer the vaccines are delayed, the later the highest peak appears since the spread of disease is not effectively controlled before the arrival of vaccines and the infectious population keeps growing. Besides the longer time needed to achieve containment, the delay of vaccine supply also significantly increases the cumulative infections. Therefore, healthcare departments and decision makers should closely monitor possible infectious disease outbreak and produce or order vaccines in the early stage of the outbreak to minimize total infections.

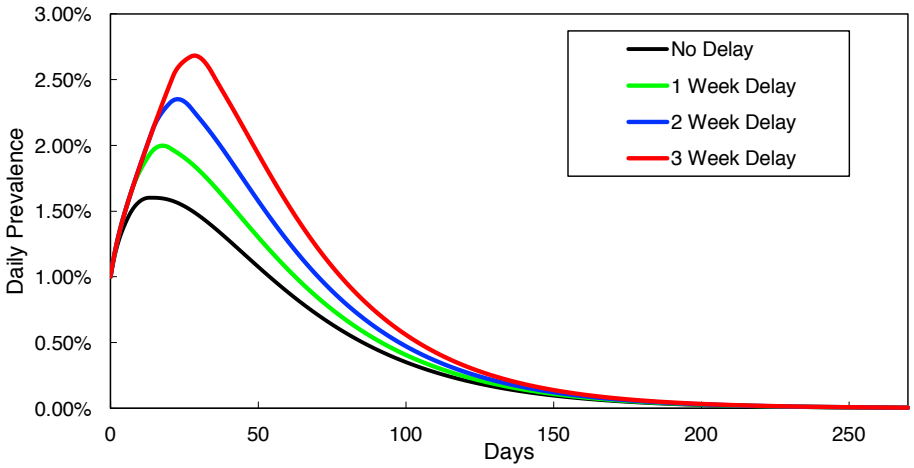


Figure 2.11: Smallpox: daily prevalence with delayed supply of vaccines.

To better understand the impact of the delay of initial vaccine supply in containing the smallpox outbreak, we contrast the overall attack rate and mortality rate of three different supply scenarios under a time of delay ranging from 0 to 3 weeks assuming optimal mixed dispensing strategy in all scenarios with $\mathcal{R}_0 = 5$ and 1% initial infection. The result is shown in Figure 2.12. With increasing initial delay, the final attack and mortality rate also increases in all three scenarios: with no initial delay of vaccine supply, the attack rates of the three supply levels are 9.32%, 6.97%, 5.35%, respectively; when the vaccines are delayed by 3 weeks, the attack rates by containment become 12.32%, 10.13%, and 8.34%, an increase of over 30% in all cases. This result again underscores the importance of the early availability of vaccines in infectious disease outbreaks and suggests that the

vaccination process should be started as early as possible.

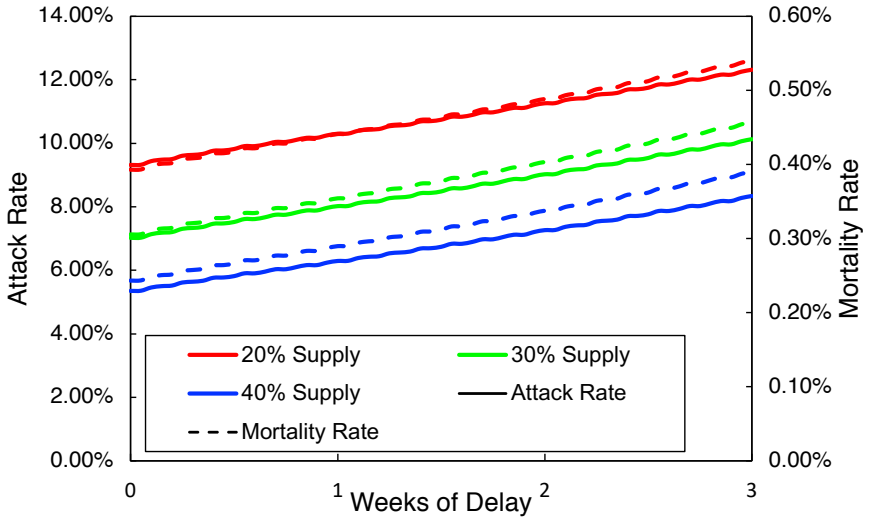


Figure 2.12: Smallpox: overall attack and mortality rate with delayed supply of vaccines.

Figure 2.13 shows the optimal switch triggers of mixed strategy corresponding to different basic reproduction number, initial percentage of infection, and the level of vaccine supply. The left panel shows that the optimal switching trigger drops below 100% in all combinations of $(\mathcal{R}_0, \alpha)$ when the vaccine supply level exceeds 18%. This conclusion can simplify the determination of optimal switch trigger in real emergencies: when the level of vaccine supply is below 18%, all of them should be dispensed to high-risk population; the switch between strategies needs to be considered only when the vaccine supply exceeds this threshold. The right panel is a magnified version of the region when the optimal switching trigger begins to decrease, and the vaccine supply levels at which the optimal switch trigger starts to drop are almost identical. This conclusion is also supported by the result displayed in Table 2.5. It shows the maximum vaccine supply level for dispensing all vaccines to the high-risk group under each combination of basic reproduction number and percentage of initial infection. It suggests that the performance of the mixed strategy is quite robust against the model parameters. This property is particularly useful in promoting the usage of mixed strategy: when combating with other types of infectious disease outbreaks, even if

the estimations of $\mathcal{R}_0$ and initial infection may not be accurate, it does not have a significant impact in determining the optimal switching triggers.

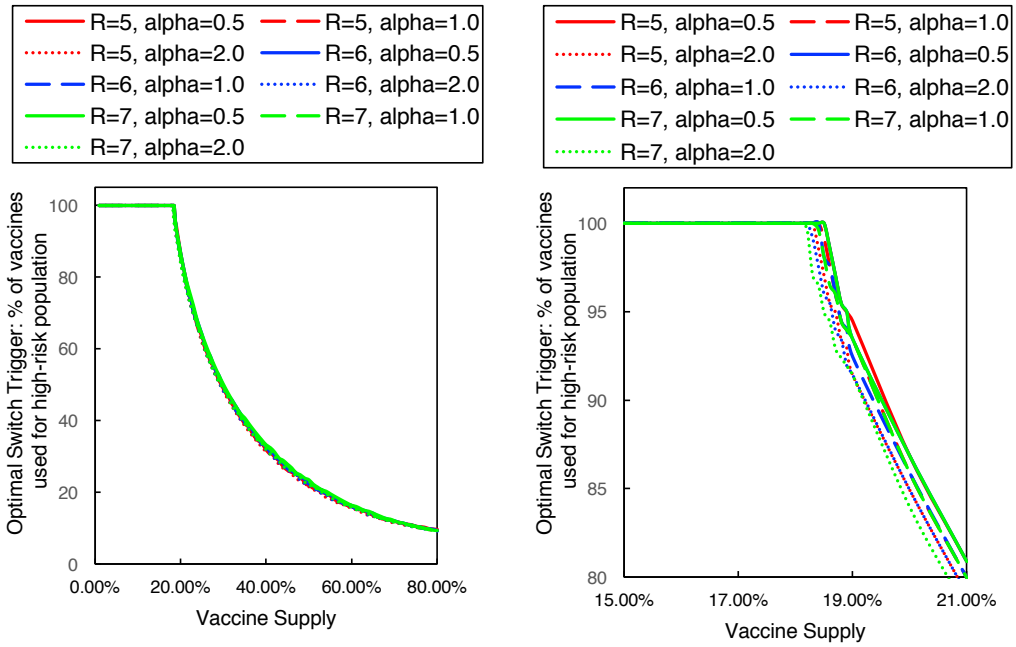


Figure 2.13: Smallpox: optimal switch triggers under different parameters and vaccine supply levels.

Table 2.5: Smallpox: maximum vaccine supply level under which $g^* = 1$

$(\mathcal{R}_0, \alpha)$	$\mathcal{R}_0 = 5$	$\mathcal{R}_0 = 6$	$\mathcal{R}_0 = 7$
$\alpha = 0.5$	18.6%	18.6%	18.6%
$\alpha = 1.0$	18.5%	18.5%	18.4%
$\alpha = 1.5$	18.4%	18.3%	18.3%

Figure 2.14 shows the attack and mortality rate at containment under three vaccine supply levels at optimal mixed strategies against the dispensing efficiency at vaccine clinics. When the throughput efficiency is higher than 60%, the impact on controlling efficiency of the outbreak is negligible; however, when the dispensing efficiency drops below 40%, the attack rate and mortality rate will increase significantly. Therefore, maintaining a high dispensing efficiency at all vaccine clinics with limited resources is crucial in containing

disease outbreaks. Managers of vaccine clinics and PODs should optimally allocate the available resource to achieve the maximal throughput possible and train the healthcare workers on a regular basis to quickly contain the disease outbreaks.

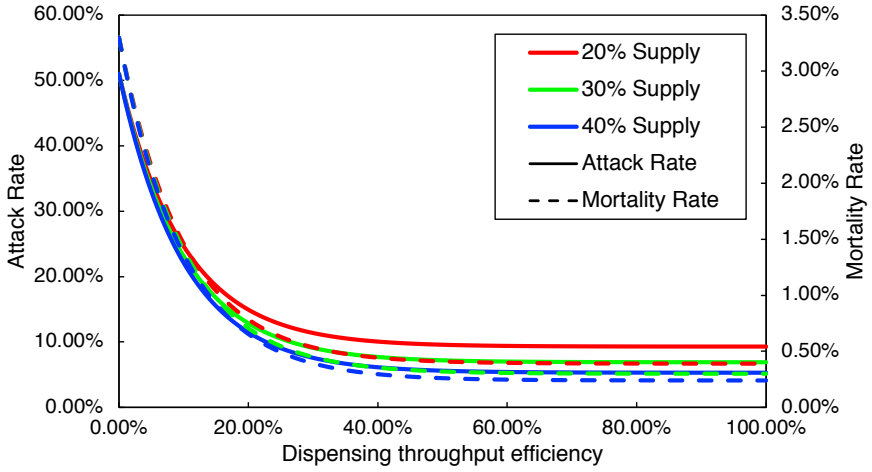


Figure 2.14: Smallpox: attack and mortality rate under different throughput efficiencies using optimal mixed strategy.

In our model, the clients visiting vaccine clinics or PODs need to be triaged and determine if they are infectious and need treatment or quarantine, or they are ready to receive vaccines. The accuracy of this triage process will also impact the trend of disease outbreak and the determination of optimal switch trigger. There are two types of errors at the triage station: Type I error (false positive), i.e., incorrectly identify the non-infectious individuals as infectious; Type II error (false negative), i.e., failure to identify the infectious individuals. Figure 2.15 shows the impact of the triage errors. The left panel shows the change of attack and mortality rates under three vaccines supply levels operated with the optimal mixed strategy against the triage accuracy with respect to Type II error. At 100% accuracy (no Type II errors), the overall attack rates are 9.32%, 6.97%, 5.35%; at 50% accuracy, the attack rates become 9.46%, 7.17%, and 5.49%. The mortality rates show similar trends. Therefore, with decreased triage accuracy regarding false negative, the number of infections and mortalities will slightly increase, as the vaccine clinics and PODs fail to quaran-

tine some infectious individuals and administrate vaccines to them, which has no protective effect. The right panel shows the optimal switch trigger with respect to the triage accuracy of Type I and Type II errors ranging from 90% to 100%. Though the accuracy with respect to false negative has an impact on the number of infections and mortalities, it does not significantly change the corresponding optimal switch trigger. On the other hand, the optimal switch trigger decreases as the triage accuracy with respect to false positive decreases (when triage accuracy is 90%, the optimal trigger for the three levels of supply is 75.27%, 41.17%, and 24.13%, as opposed to 86.85%, 48.25%, and 29.92% when the accuracy is 100%). Since more healthy/susceptible individuals are triaged as infectious and send to hospitalization without dispensing vaccines, fewer vaccines will be dispensed, and the time of switching to non-prioritized strategy will become earlier.

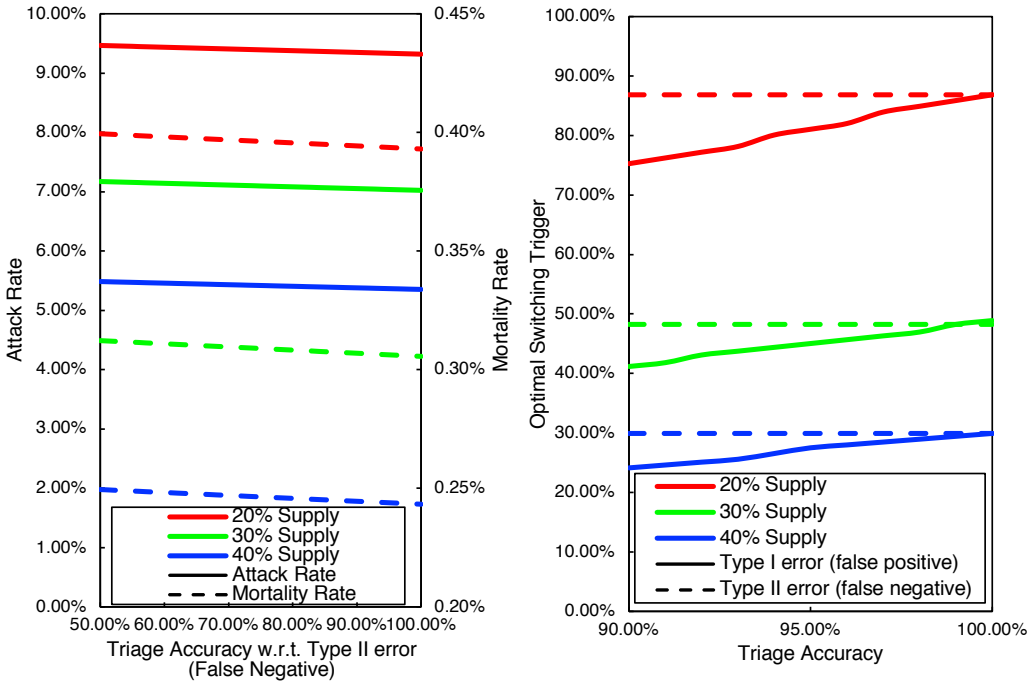


Figure 2.15: Smallpox: effect of triage accuracy at the PODs/vaccine clinics.

Discussion

In this study, we propose a generic modeling framework for infectious disease outbreaks with human interventions. We focus on the scenario of smallpox outbreaks and study the effect of different vaccination strategies due to the limited supply of vaccine on the overall attack and mortality rate. We propose the idea of optimal switch trigger, which represents the percentage of high-risk population been vaccinated before switching to a non-prioritized vaccination strategy. We incorporate both the compartmental model for disease propagation and vaccine queuing model and develop numerical optimization on the framework. To ensure the credibility and realism, we use the data from our past time-motion study to populate the model parameters and run the simulation for the population in the state of Georgia. Sensitivity analyses are included to investigate the impact of various perturbation in implementing the optimal mixed vaccination strategy, including the delay of vaccine supply and inaccuracy in triage.

Our analyses suggest that the optimal mixed vaccination strategy outperforms the full-prioritized or full-non-prioritized strategies in terms of minimizing overall attack and mortality rate under different combinations of model parameters and availability of vaccines. When the vaccine supply is limited, all vaccines should be dispensed to high-risk groups for best containment results; when the vaccine supply exceeds a threshold, a switch in the dispensing strategy can be favorable and results in less attack and mortality rate. Delay of availability of vaccines can have a significant impact on the trend of the disease outbreak and may result in an increase in cumulative infectious population, thus early intervention and availability of vaccines are crucial in containing disease outbreaks. The result of implementing the optimal mixed strategy can also be influenced by various factors, including the operation efficiency at vaccine clinics/PODs, and the triage accuracy. Though the optimal switch trigger is robust to some of the parameters, public health departments still need to ensure the efficiency of vaccine dispensing and the accuracy of triage via the usage of

electronic health records (EHR), electronic scanning technologies, and optimal design of POD layout and allocation of healthcare workers to achieve the best containment result.

This modeling framework is highly flexible: it is not constrained to a single type of disease or intervention method. To model different types of diseases, e.g., vector-borne diseases like dengue and yellow fever, or food-borne disease like cholera, decision makers need just to plug in the corresponding structure for the compartmental model and estimations of parameters based on the biological facts, and use the modeling framework and computational engine to derive the optimal switch trigger for the best mixed vaccination strategy. In addition, our definition of optimal switch trigger integrates not only the property of disease propagation itself but the vaccine supply and operations in the actual vaccine dispensing operations as well. Other important factors impact the decision-making process are also highlighted in the sensitivity analysis, making it a practical and effective tool to use in real emergencies.

With increasing danger of exposing to disease outbreaks worldwide and insufficient supply or even unavailability of vaccines, it remains a challenge to determine the best way to protect more people with limited resources. Instead of focusing on a microscopic view of immunology effect prediction on the individual level for optimal utilization of vaccines, we tackle this problem in a macroscopic view of analyzing the strategies for dispensing. One advantage of this approach is that the result is almost independent of how the vaccines work; decision-makers just need to know the overall vaccine effectiveness to derive the optimal switching trigger. Also, sensitivity analyses indicate that the optimal switch trigger is highly robust against multiple parameters, thus a near-optimal result can be easily achieved even when accurate estimations of model parameters are not immediately available. This makes our modeling framework more attractive in practical usages and enables governments and healthcare departments to make a rapid reaction in case of emergencies.

Though we use smallpox outbreak as an example to demonstrate the usage of our modeling framework, the conclusion can easily be generalized to contain different types of infectious disease outbreaks. For example, we apply a similar modeling strategy to optimize the medical resource allocation in the Ebola virus outbreak in West Africa during 2014 to 2015; the similar approach was also applied to study the containment strategy of the Zika virus outbreak in Brazil and Puerto Rico. The current yellow fever outbreak in Brazil, Angola, Uganda, and DR Congo is proven difficult to control due to various reasons: changing the climate, insufficient insect repellent or pesticide to control the mosquito population, and insufficient supply of vaccines. Currently, there is only one manufacturer of the yellow fever vaccine in the U.S., and the shortage of vaccines is likely to continue until mid-2018 due to the delay of manufacturing [117, 151]. This signifies the importance of our work, as it provides a way to maximize the effect of protection using available vaccines. The generic model for the vector-borne disease can be plugged into the generic modeling framework with intervention and provides the optimal switch trigger given the current supply level of vaccines and the proper segmentation of risk groups. The resulting strategy and policy may have a major impact on the trend of the current yellow fever outbreak, and protect population health in a broader region.

CHAPTER 3

DECISION SUPPORT SYSTEMS FOR BIOSURVEILLANCE AND OPERATIONS

Biosurveillance and emergency medical operations are crucial during public health emergencies. Public health administrators need to understand the current situation of the emergency, check the availability and inventory of resources to mitigate emergencies, and evaluate different intervention strategies and choose the most effective one to implement. Due to the nature of public health emergencies, these decisions need to be made in real time with minimal errors. Computerized decision support systems and information technologies are reliable and efficient for these applications. They are widely used in multiple public health applications to facilitate and enhance the decision-making process for better mitigation results and population protection. In this chapter, we discuss a decision support system suite for biosurveillance and operations: the RealOpt suite. We focus on one module in the suite, RealOpt-Regional, and discuss its system design and applications.

RealOpt is an interactive web-based real-time decision support suite. The system integrates visualization, information and cognitive analytics, and dynamic large-scale computational modeling and optimization tools that allow public health emergency preparedness coordinators to determine optimal response facilities and locations, resource needs and supply routes, and population flow in real time. With an eye towards flexibility and future system expansion, RealOpt is designed in modular format allowing direct linkage to multiple functional modules. Currently, the system has twelve modules covering emergency response preparedness and operations for biological, chemical, radiological/nuclear incidents, biosurveillance, epidemiology, and decontamination models, operations logistics and networks, a real-time crowd sourcing data feed, and evacuation planning. RealOpt has been used for biodefense and H1N1 regional planning and operations, regional flood and

hurricane responses, 2010 Haiti earthquake disaster relief, 2011 Japan Fukushima disaster, 2014-2015 Ebola containment assistance and after-event public health preparedness training in West Africa, and current Zika virus containment analysis. The fast solution engines enable real-time use for rapid decision and scenario analysis, since it requires only one CPU minute to determine an optimal network of facilities and resource needs to serve a population of over 10 million.

3.1 Introduction

Planning for a catastrophe involving a disease outbreak or mass casualties [27] is an ongoing challenge for first responders and emergency managers. They must make critical decisions on treatment distribution points and staffing levels, and gauge the potential impact on affected populations in a compressed window of time when seconds could mean life or death. Although extensive efforts have been made to plan for a worst-case scenario on the local, regional and national scale, the U.S. Government Accountability Office (GAO) found gaps still exist [99]. While many states have made progress in planning for mass casualty events, many have noted concerns related to maintaining adequate staffing levels and accessing other resources necessary to effectively respond.

After outbreaks of potential public health emergencies, immediate and aggressive responses must be carried out. This includes medical countermeasures dispensing [118, 208] to mitigate possible deaths and to control epidemic. Besides dispensing medication, "service constructs" for food and water distribution, temporary shelters, medical care, screening and registry, and decontamination may also be required. Without loss of generality, we call such sites point-of-dispensing sites (PODs). These are the places where affected individuals come to receive services.

While dispensing medical countermeasures and providing medical services require spe-

cific healthcare personnel, distribution of food and water and other personal needs can be carried out by other emergency workers and volunteers. Nonetheless, both types of services share key elements: resources are scare, time is precious, risk is uncertain and evolving, there is a large affected population to serve, and the on-the-ground conditions can be exceedingly stressful, for the impacted population and for the emergency workers. To maximize the throughput that can be served under limited resources of labor, time, and potential damaged infrastructure, PODs must be established such that they are flexible, scalable, sustainable, and agile for operations continuity and fluidity. Policy and decision makers must determine where to establish a network of sites to achieve the maximum overall benefit. Further, POD layout must be designed to facilitate the best operations workflow and throughput while minimizing resource requirements.

CDC's Cities Readiness Initiative (CRI) is designed to enhance preparedness in the nation's largest population centers to effectively respond to large-scale public health emergencies needing life-saving medications and medical supplies. State and large metropolitan public health departments develop, test, and maintain plans to quickly receive medical countermeasures from CDC's Strategic National Stockpile and distribute them to local communities. RealOpt is designed and used by many such departments to assist in these efforts.

For a potential anthrax response, the goal of dispensing countermeasures to the affected population within 48 hours of dispensing decision presents real challenges. A commercial system that offers a discrete event simulation tool for scenario analysis, but takes over 10 hours to generate a feasible solution per scenario, is not suitable for real-time planning. In contrast, scenario analysis and resource optimization in RealOpt is typically completed under one minute, so is well-suited for real-time planning.

RealOpt is an interactive online software enterprise for large-scale regional medical dispensing and emergency preparedness and responses. It features interactive visualization tools to assist with spatial understanding of important landmarks in the region, assess the population densities and demographic makeup of the region, and identify potential facility locations. It also features backend mathematical models for large-scale facility location, resource allocation optimization, and dynamic disease propagation problems. The system is equipped with novel rapid solution engines for strategic and operational planning, real-time dynamic optimization, and epidemiological analysis.

In the following sections, we first describe the system architecture and design of RealOpt. Next, the system modules are presented, beginning with functionalities that relate to the user experience, and moving on to functionalities that pertain to managing geographical boundaries; facility location and population assignment; ZIP code and population composition; multi-modality dispensing; POD layout design and resource allocation; disease propagation analysis and biosurveillance; and regional information sharing, reverse reporting, tracking and monitoring, and resupply. Application of the system to planning for mass dispensing for anthrax response, and to planning and responding to actual disasters will be briefly presented. The section concludes with a discussion of future work and challenges.

3.2 System Architecture and Design

With flexibility and expansion in mind, RealOpt is designed in modular form (Figure 3.1) and allows direct linkage to functional modules. Currently, it is composed of twelve primary modules for integrating system interactivity, client-server architecture, optimization, and policy/decision support. The foundation web-component in the systems design is the RealOpt-Regional module.

RealOpt-Regional is implemented via the client-server architecture. For clients, the

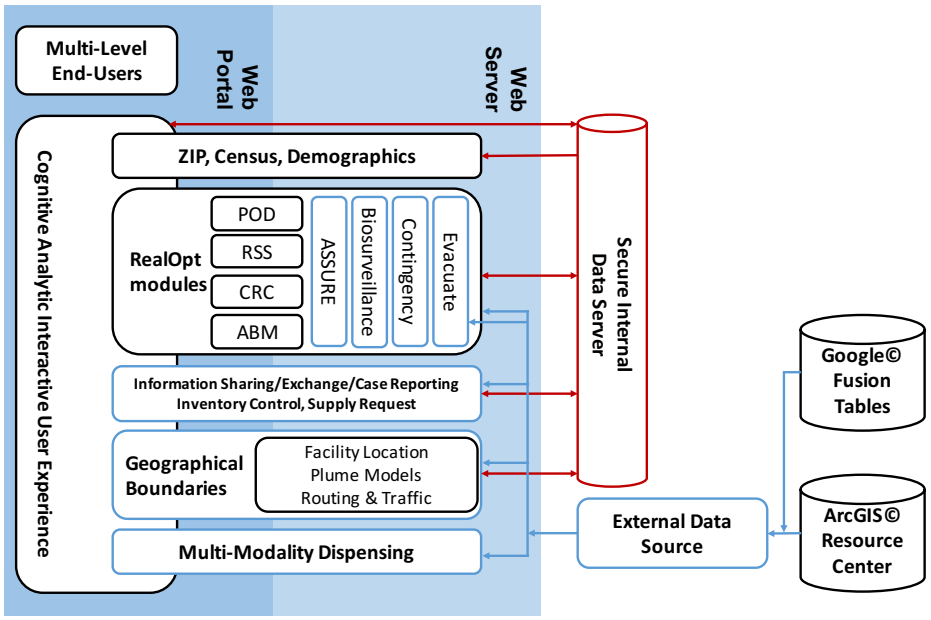


Figure 3.1: RealOpt suite: system modules.

Google Maps API along with a webpage are incorporated to provide the primary I/O functionalities. On the server, a rapid hybrid large-scale algorithm that couples features of a combinatorial genetic algorithm and an adaptive greedy approach is implemented as a solution engine for optimization of the network of facility sites and population flow and assignment. The intensive optimization computation is performed on the server, and requires no extra computing effort from the client. The user's system is used primarily for webpage display and object rendering on the map.

Asynchronous JavaScript and XML (AJAX) with basic access authentication is used for sending and receiving requests asynchronously. In RealOpt-Regional, a large-scale facility location problem may take seconds or a minute to initiate and solve. AJAX enables sending the optimization request from the client to the server wherein users do not have to wait for response from the server during optimization. Users at this point are free to explore other features in RealOpt-Regional. Further, even if the client becomes disconnected from the server, the user can still retrieve the optimization results. Potential internet disruption is unavoidable; hence this feature is very appealing and important. While the connection

can be manually or automatically re-built, the running progress and solutions will not be accessible unless another request is sent from client web browser to the server. AJAX enables users to obtain updated running progress or optimization results once the program runs to completion.

Java Native Interface (JNI) is employed to link multiple modules on the server to complete the backend computation and interface. The system architecture is designed so that only the junction module needs modification when updating the solution engine. A schematic design of the RealOpt-Regional system architecture is depicted in Figure 3.2.

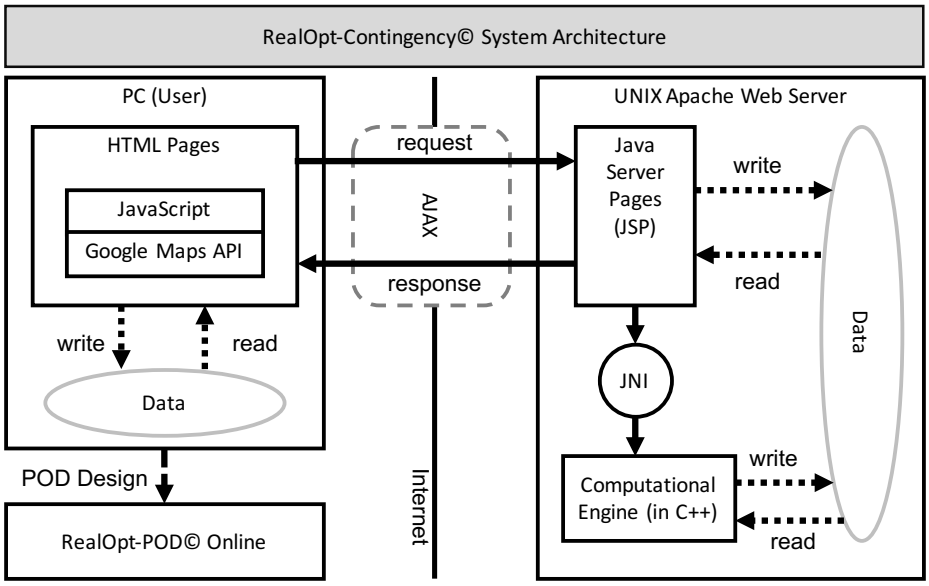


Figure 3.2: RealOpt-Regional: schematic design.

3.3 System Modules and Functionalities

We briefly describe some of the RealOpt system modules and functionalities below.

3.3.1 Interactive User Experience

We incorporate the graphical interactivity enabled by Google Maps JavaScript API as part of the I/O functionalities. With up to 100,000 map loads per 24 hours, Google Maps JavaScript API provides an efficient and stable mapping environment for users. This module equips users with spatial understanding of important landmarks in the planning and surrounding regions. Also, without the need for prerequisite knowledge in mathematical programming, users are empowered to build, optimize, and evaluate service networks in a graphical visualization environment. They can interact with the map by selecting jurisdictions, adding/removing potential PODs (as markers), defining dispensing constraints (e.g., capacity constraints in each facility, and travel distance/time constraints), and highlighting population served, household characteristics and various parameters (e.g., average number of members per family; ethnicity/language represented in the population). In the backend, the associated mathematical model will be built automatically and sent to the server for optimization.

Other capabilities are implemented to allow users to manage their files on the secured server directly from their web browser (e.g. adding, saving, uploading, and removing). This creates a user experience like that of using a PC-application even though it is in the web environment. The interactivity is designed to be compatible with mobile tablet devices. Figure 3.3 shows the interface of RealOpt-Regional for public health emergency preparedness coordinators and planners. On the left is a Google map, and on the right are six function panels: “Jurisdiction”, “POD (Marker)”, “Optimization” (Facility Location), “DrawPolygon”, “Multimodes”, and “Twitter”. Help files and step-by-step practice scenarios are provided to guide users in building their own planning cases and to familiarize them with usage of the system.

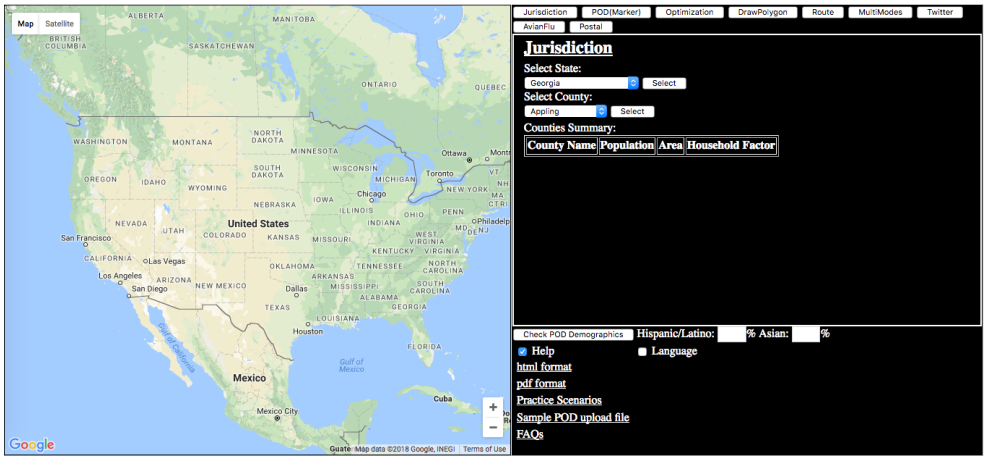


Figure 3.3: RealOpt-Regional: main planning interface.

3.3.2 Geographical Boundaries

Through ArcGIS, RealOpt-Regional can display the geographical county boundaries of the United States. Emergency response planners can select and define their planning area spatially. Demographics such as population and density of each county are also provided to complete the necessary input when planners are building and optimizing the dispensing/service networks. City boundaries can be defined by free drawing polygon tools "DrawPolygon" on the map. This enables RealOpt-Regional to work compatibly with the Cities Readiness Initiative to enhance preparedness of our nation's largest cities and metropolitan statistical areas. The flexibility of users manually drawing planning areas also ensures that the heterogeneity of population densities between cities and counties can be considered in building the dispensing networks. Further, such a free drawing feature is critical for international sites (e.g., the disaster response effort for the 2010 Haiti earthquake) where populations are on the move, and affected regions are dynamic. The drawing process, which includes tools for adding/deleting a vertex, and removing and modifying the polygon shape, simply amount to a series of clicking and dragging operations on the map. Users can save their drawings for future use. Figure 3.4 illustrates an example use of "DrawPolygon" in outlying the boundary of the city of Chicago.

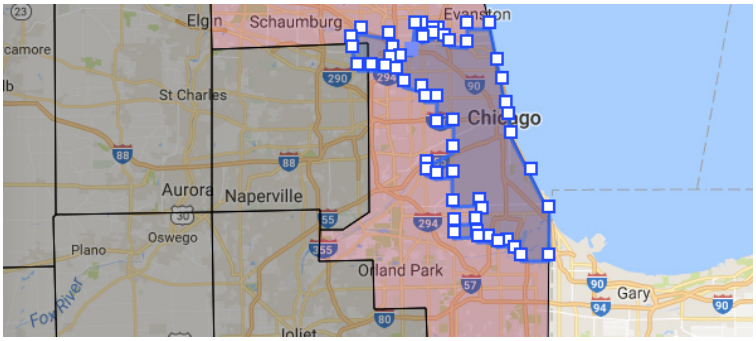


Figure 3.4: RealOpt-Regional: user-defined boundary of the city of Chicago.

3.3.3 Network of Service, Locations and Population Flow and Assignment

Once a region is selected or drawn, users can input potential PODs by loading a file with addresses, input the addresses interactively, or simply drag-and-drop the POD markers onto the region. Physical constraints on the facility must be modeled, e.g., capacity of a facility cannot be violated (e.g., parking capacity, fire codes, and building regulations limit the total number of individuals that can be present simultaneously inside a facility). Users can specify this globally for all PODs, or more commonly, they can specify the value for each POD. Travel distance, household factors, and time for completing services at the facilities can also be input.

Various objectives are incorporated within the RealOpt-Regional computational engines. In the event of catastrophic incidents, it is critical that PODs are strategically located to allow easy access by the affected public. Hence, minimizing transportation time and distance is one critical objective. Further, the setup and operating costs of PODs cannot be neglected. In our model, setup costs and travel time form the composite objective criteria.

A POD must be accessible by service workers. It should include a good communication infrastructure, and must be readily protected by law-enforcement personnel. The facility must be capable of handling a large influx of people with good access routes for supply and

resupply. The spatial maps allow the users to evaluate the surrounding POD areas to ensure that candidate POD selections are deemed appropriate.

The planning area may include a set of counties and/or a composition of a free drawing region. By default, it is discretized at 40-by-40 resolution. The population within the planning area is assumed to be distributed evenly, or it can be strategically specified. Once the set of potential PODs are selected with desired parameters, RealOpt-Regional translates this input into a facility location problem on the backend. Specifically, given a selected jurisdiction and a regional population, under the capacity and travel distance constraints, setup cost and travel distance/time as objectives, and the required completion time for protecting the chosen regional population, we must determine a) an optimal set of PODs and their respective locations needed for cost-effective operations; and b) the population assignment to this optimal set of PODs.

The computational engine consists of a two-stage facility location integer program. The first stage optimizes the number of facilities opened. This is formulated as a capacitated POD-location problem COVER-CAP to ensure that at least two PODs are opened. If a catastrophic event at one site necessitates shutting down a POD, the remaining location can continue to carry out the emergency dispensing. The model ensures that each household will travel at most the maximum allowed travel distance and time, every household is served, and that the capacity of the facility is not violated. COVER-CAP returns the minimum total number of PODs needed. Next, the MINAVG-CAP problem is formulated to minimize the travel distance and time over all households while keeping the number of PODs fixed to the optimal value that the COVER-CAP problem determined. See [136] and [142] for variations of the full IP formulations.

For a jurisdiction of about 20,000 households, current competitive commercial opti-

mizers fail to solve any of the instances after running for several months of CPU time. The difficulty is partly because facility location remains an *NP-hard* problem and partly due to the diverse requirements of the facilities and the very mixed population densities. Using recent computational advances for solving intractable facility-like instances [144], we were able to solve instances ranging from 200,000 to 10 million variables within 40,000 CPU seconds. To facilitate strategic and operational planning and real-time dynamic decision analysis, we design a novel rapid solution engine that couples a combinatorial genetic algorithm with an adaptive greedy search [141]. The optimization heuristic schema is shown in Figure 3.5. The engine is implemented in C++. Algorithmically, users can wait until the optimization is completed (usually takes only seconds), or they can interrupt the solution process manually to obtain intermediate non-optimal solutions.

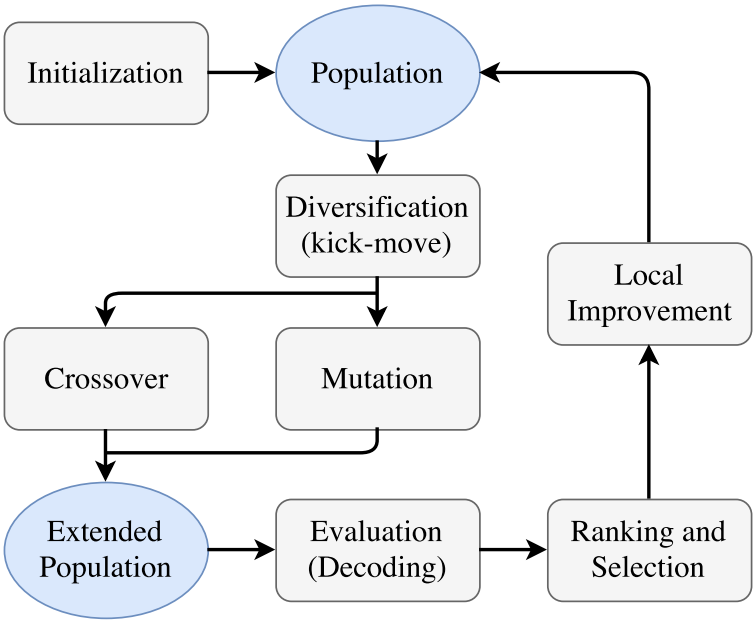


Figure 3.5: RealOpt-Regional: algorithmic flow of our facility location heuristic.

3.3.4 ZIP Code and Population Composition

By rearranging and analyzing ZIP code demographics and area boundaries, RealOpt-Regional maintains a set of up-to-date census data bank that contains 32,036 postal-code

areas in the United States. The system provides ZIP code based information including regional demographics, social indicators, economic indicators, and boundary of ZIP code areas.

After determining a set of optimal POD locations and population assignment, users can query POD coverage based on demographics by overlaying the POD service area with all adjacent ZIP code areas, and aggregating ZIP code based data by considering size of overlapping, and population density of each overlapped ZIP code area. POD coverage based demographics are dynamic and dependent on user-defined parameters in optimizing the locations and the total number of facilities. Using this information, emergency planners can identify appropriate personnel for each specific POD and/or mode of dispensing (e.g. pediatric assistants, translators, transportation needs or mobile PODs for socially-disadvantaged individuals). Figure 3.6 illustrates the population composition by overlapping POD coverage with all adjacent ZIP code areas. We can identify demographics of the majority population served by the POD.

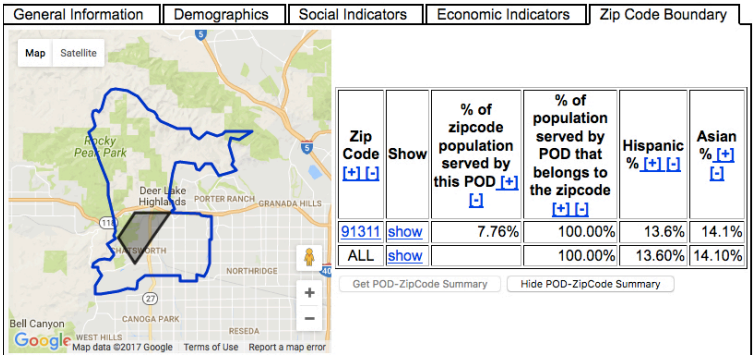


Figure 3.6: RealOpt-Regional: POD demographics for 91311.

3.3.5 Multi-Modality Dispensing & Public-Private Partnership

The goal of mass dispensing is to protect the worried well (general population) efficiently and effectively under time constraints. In an anthrax attack, it is important that the affected

population receive antibiotic prophylaxis within 48 hours of the determination that an attack has occurred, since the mortality rate for persons demonstrating symptoms of inhalation anthrax is extremely high. Multiple dispensing modalities often must be employed to serve the entire regional population rapidly [136, 142].

Open PODs (i.e., open to the public) can be drive-through or facility-based walk-through. Special needs populations may need closed PODs on site. For example, nursing homes, assisted living facilities, homeless shelters, hospitals, and prisons house many residents for whom it would be inconvenient or inadvisable to travel to a public dispensing facility.

Further, self-organized closed PODs should be promoted. Corporate offices with many employees could protect their own. Once these sites receive prophylactic supplies, they could set up a closed POD with their own healthcare staff and volunteers, or with public health staff supplemented by the state. Several factors suggest that such closed PODs will have fewer security concerns and will be easier to manage than public PODs. These factors include familiarity with the environment and people (e.g., fellow residents/employees), existing security measures including established checkpoints and previously authenticated identification badges with photo and/or biometric markers, and less stress than having to commute to a public POD.

Airports and hotels where many non-resident travelers can be found are also candidates for setting up PODs. This is particularly important with infectious disease outbreak to ensure that travelers are protected with medical countermeasures before they leave for elsewhere. Universities can use their own health facilities (and if necessary, additional mobile on-campus PODs provided by the state) to provide prophylaxis to on-campus students, staff and faculty. Clearly, if large employers and medical facilities provide prophylaxis

to their own employees, families, and patients, it will eliminate a high percentage of the population (may be as high as 40% in some large cities) from visiting public PODs, thus reducing the burden on these facilities.

POD markers in RealOpt-Regional can be marked as closed or open, drive-through or walk-through, and with specific demographics needs. Currently the system maintains a data bank of 2,000 homeless shelters, 4,200 universities and colleges, 5,128 airports, 12,881 high schools, 73,079 assisted living facilities, and 24,192 hotels. Public health and emergency response planners are equipped with this information and can make better decisions in constructing more cost-effective and efficient mass dispensing networks. Users can also search specific business or facility types to explore options of dispensing sites.

3.3.6 POD Layout Design and Resource Allocation

RealOpt-POD [138, 139] was the first RealOpt module. It was designed and implemented in 2003, and has been advanced through the years. It currently has over 10,000 public health site users. Its users also include fire departments, police departments, and hospital networks. Written entirely in Java, it is a stand-alone computerized decision support system for facility layout, resource allocation, and intra-facility disease propagation analysis. The real-time solution capability is supported by automatic graph drawing tools, and our in-house large-scale simulator and optimizers. Since 2009, a branch version has been developed within the RealOpt web-based client-server architecture. It is also interfaced with RealOpt-Regional. While the automatic graph-drawing tool and user interface are built into applets and embedded in the webpage, just as RealOpt-Regional, the core computation for optimization and large-scale simulation are powered by the server.

The linkage of RealOpt-POD and RealOpt-Regional allows users to design POD layouts and perform resource allocation optimization on the spot once the network of facilities

is determined. However, the standalone version RealOpt-POD is still in high demand since planners do not need internet access to use it.

3.3.7 Radiological Module

RealOpt-CRC is designed in parallel to RealOpt-POD for population screening in response to radiological/nuclear incidents. There are major overlaps in these two systems wherein CRC incorporates the decontamination, bioassay and other processes that are critical for radiological screening and population registry and protection. It was used in response to the 2011 Fukushima Daiichi radiological disasters (See Section 3.4).

3.3.8 Biosurveillance

Large-scale dispensing of medical countermeasures has proven to be an effective way to contain outbreaks of highly infectious diseases, such as smallpox, pneumonic plague, and pandemic flu [118, 208]. Further, successful dispensing operations, such as optimizing POD operations through RealOpt-POD, can help reduce resource requirements, shorten patient waiting time and queues, and maximize throughput [139, 141]. However, the large influx of individuals brought into these dispensing centers raises the potential risk of intra-facility cross-infections. RealOpt supports surveillance of biological attacks or natural pandemics on three fronts:

* Intra-infectivity tracking. In the planning phase of a dispensing event, POD planners and epidemiologists can employ RealOpt-POD for estimating the dynamics of intra-POD infection under different scenarios of POD design and personnel allocation. The epidemiology module integrates agent-based simulation within the discrete event simulation platform, offering a powerful means to track an individual's health status within the dispensing operations;

* Case incidents tracking. In addition to providing the epidemiological estimates during the planning stage, POD managers can report and register incidents via RealOpt

in real time during the dispensing event. These numbers, along with the attached timestamps, can provide valuable information in monitoring and containing the disease spread. They are also beneficial for epidemiologists to collect data and refine model parameters in real time;

* Incorporation of plume models within regional planning. RealOpt-Regional allows automatic import of plume models and biological-infected regional zones within the planning region. Users can also import plume radius and coverage via the free drawing panel. This information can be critical for population protection planning, allowing planners to select proper/safe candidate sites (for shelters, medical service, and/or dispensing) and estimate resources required for regional coverage.

Beyond strategic and operational planning, training and cost-effectiveness analysis, RealOpt also offers various information sharing/exchange and tracking capabilities. Specifically, it allows users to:

* Input medical supply burn rates. The data will be registered with a timestamp. This allows regional commanders to oversee the dispensing process across the affected region, re-direct populations as necessary, and order re-supplies appropriately;

* Perform real-time POD reconfiguration using current service performance information through the service distribution estimator in RealOpt-POD;

* Report case incidents for disease monitoring and derivation of mitigation strategies;

* Perform reverse information sharing by highlighting traffic and road conditions in surrounding areas to alert regional managers of necessary steps to re-route transportation or re-direct citizens;

* Input plume information to alert populations of potential risks and guide them to stay outside the plume zone(s).

The information can be used by multiple stakeholders at different levels. Regional commanders are empowered with multi-level ground knowledge and performance measurements that will facilitate expediting operations, mobilizing resources, and containing threats and spread of disease.

3.3.9 Multi-Level End-Users Access

RealOpt can be used by both emergency preparedness personnel and the public to obtain service site information. Multiple levels of authorization are used to control proper access. In general, emergency preparedness personnel can:

* Determine a network of optimal service/dispensing locations and the optimal combination of dispensing modalities (e.g. combinations of drive-through, walk-through, mobile, open and closed PODs) for regional preparedness;

* Review demographics at each site and identify the appropriate personnel (e.g. translators, pediatric assistants) needed for effective dispensing;

* Determine in real-time a resupply schedule for medical countermeasures across the network through real-time feeds of product usage;

* Determine the closest facility and feasible traffic routes (incorporating inverse road conditions) to resupply, and to transport sick individuals from PODs to respective hospitals for treatment;

* Identify region-wide budget and labor needs to accommodate an emergency, and determine recruitment needs of healthcare workers and volunteers;

* Perform economic analysis on multi-modality variations.

Emergency preparedness personnel can be further categorized into three levels: “emergency planner”, “RSS (receive-sort-staging) manager”, and “regional commander”. Emer-

gency planners are provided with general functionalities mentioned above; RSS managers can model RSS sites on the map, and can provide supply when requested. Regional commanders can further manage inventory at each site, review inventory summary over all sites, review item information summary, and manage shipments whenever necessary. Individual POD managers can login to design POD layout and optimize POD performance via RealOpt-POD. They can register/report burn rates, case incidents, and request resupply. Epidemiologists work along POD and Regional commanders to analyze disease spread and design mitigation strategies.

The system will be open to the public only during an actual event to obtain driving directions to primary and backup PODs that are specifically assigned to them. They can also review relevant information related to the specific medical countermeasures dispensed, and to register for the medication.

3.4 Applications

RealOpt has been used for planning for mass dispensing for anthrax response, actual vaccination events for seasonal flu, 2009/2010 H1N1 response, and hepatitis B vaccination. The system has also been tailored for the 2010 Haiti earthquake emergency response, the 2011 Japan Tohoku-Fukushima radiological disaster efforts, the 2014 Ebola response and training operations, and the current Zika virus strategic containment analysis.

3.4.1 Biodefense Mass Dispensing Regional Planning

To illustrate the strategic capability and importance of RealOpt, we briefly describe the biodefense planning and capability and resource requirement estimates for New York City. The city consists of five counties: Bronx County (Bronx), Kings County (Brooklyn), New York County (Manhattan), Queens County (Queens), and Richmond County (Staten Island). According to the U.S Census Bureau, the total population in New York City was

estimated to be 8,550,405 in 2015. The land area was 303 square miles, and the number of persons per square mile is 27,603. Among the entire population, 47.60% speak languages other than English at home.

For brevity, we illustrate only one analysis here where the maximum travel distance is set to 15 miles, and the physical capacity of each facility ranges from 1,000 to 2,000 per hour, reflecting the actual site specification. Two 12-hour shifts are operating at each location to ensure that the population is protected within 24 hours. Using census data, it is assumed that an average household has 2.57 people.

Figure 3.7 shows a POD network returned by RealOpt-Regional that covers 96.3% of the New York City population. A total of 101 PODs are needed, each runs two 12-hour shifts. Figure 3.8 illustrates the percentage of non-English speaking (Hispanic/Latino and Asian) population served by one of these POD locations. Using this information, emergency planners can contrast the type and number of interpreters needed to execute the necessary POD operations.

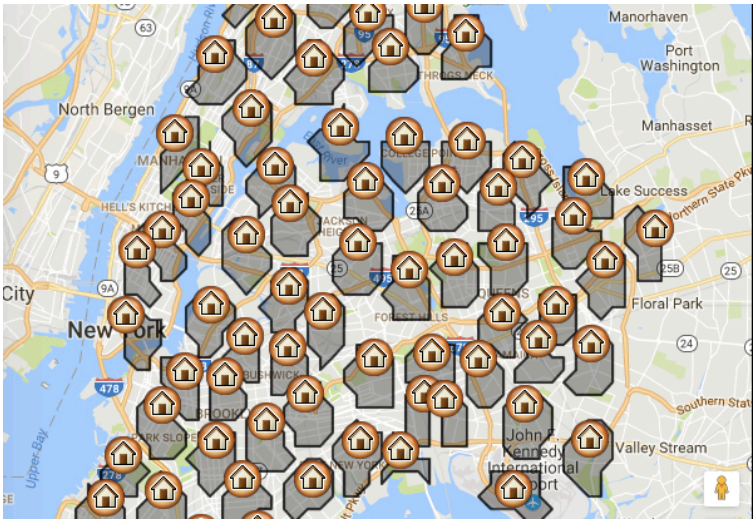
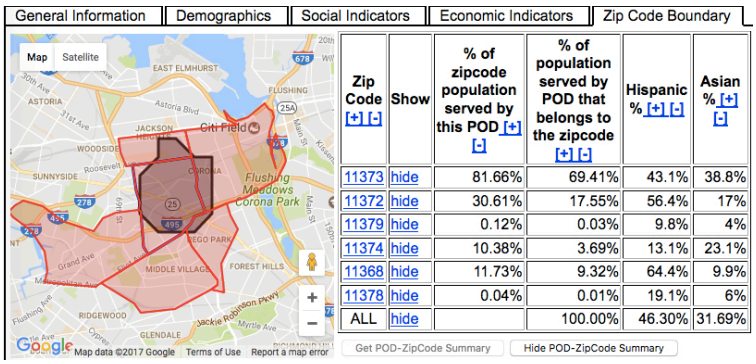


Figure 3.7: RealOpt-Regional: Optimal POD facility network for New York City.



Zip Code [+] [-]	Show	% of zipcode population served by this POD [+] [-]	% of population served by POD that belongs to the zipcode [+] [-]	Hispanic % [+] [-]	Asian % [+] [-]
11373	hide	81.66%	69.41%	43.1%	38.8%
11372	hide	30.61%	17.55%	56.4%	17%
11379	hide	0.12%	0.03%	9.8%	4%
11374	hide	10.38%	3.69%	13.1%	23.1%
11368	hide	11.73%	9.32%	64.4%	9.9%
11378	hide	0.04%	0.01%	19.1%	6%
ALL	hide		100.00%	46.30%	31.69%

Figure 3.8: RealOpt-Regional: ZIP code-based demographical information.

Next, RealOpt-POD is used to design POD floorplan and predict the optimal resource allocation, based on the POD operations service distributions collected from local jurisdiction. Figure 3.9 shows the number and types of workers needed to cover 96.3% of New York City population with or without splitting the PODs into Hispanic-operated versus non-Hispanic operated. The number of interpreters in the splitting case is significantly smaller than the non-splitting case. This indicates that the PODs in the splitting case are easier to set up (for just-in-time training and managing) for efficient operation during emergencies.

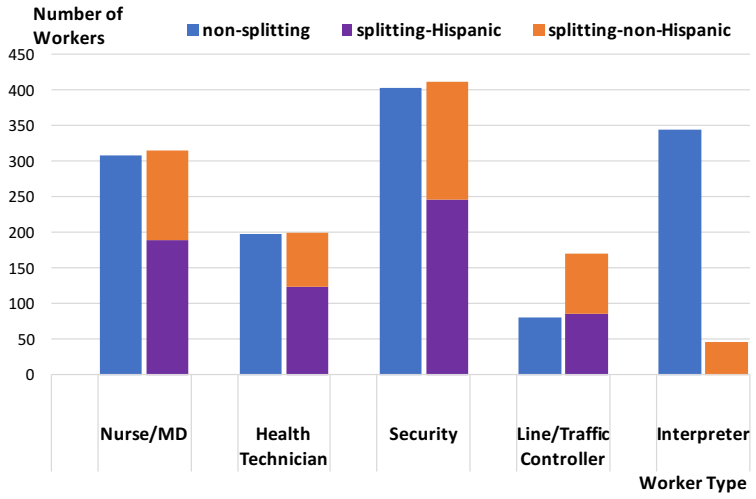
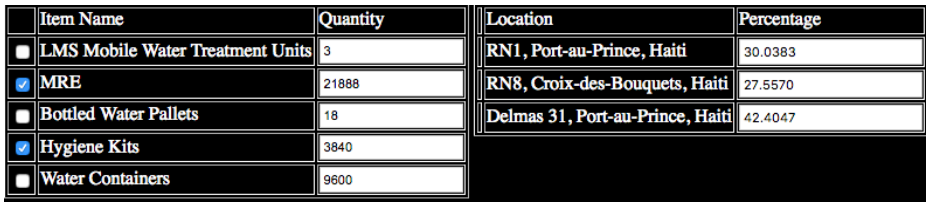


Figure 3.9: RealOpt-Regional: Number and types of workers needed to cover 96.3% of New York City population.

3.4.2 Real-Life Disaster Response Effort

Large-scale disaster and humanitarian relief efforts (e.g., in response to earthquakes, hurricanes, forest fires) where homes are destroyed, critical infrastructures are damaged, and tens of thousands or millions of people's lives are affected, require rapid establishment of "service constructs". These service constructs serve as camp shelters for the displaced population; as distribution nodes for receiving supplies for on-the-ground responders; as dispensing sites for handing out food and water to the affected population; and as hospital tents for medical care of the injured and sick.

In 2010, RealOpt was tailored to RealOpt-Haiti for part of Haiti's emergency response efforts. In this case, distribution nodes were established, with affected populations estimated by region based on on-the-ground dynamics and movement of the population. Figure 3.10 shows the supplies arriving at the airport and seaport, and the population density across the 11 in-land distribution nodes. The system allows distribution optimization and provides the responders the items and quantities of shipment to each distribution node. Figure 3.11 shows the distribution paths and supply items and quantities from the supply nodes (airport and seaport) to the in-land distribution nodes returned by RealOpt-Haiti. The system was also used in optimizing resource allocation and operations within each distribution node, temporary medical tents, and camps/shelters for the displaced population.



	Item Name	Quantity
☐	LMS Mobile Water Treatment Units	3
☑	MRE	21888
☐	Bottled Water Pallets	18
☑	Hygiene Kits	3840
☐	Water Containers	9600

Location	Percentage
RN1, Port-au-Prince, Haiti	30.0383
RN8, Croix-des-Bouquets, Haiti	27.5570
Delmas 31, Port-au-Prince, Haiti	42.4047

Figure 3.10: RealOpt-Regional: supplies at airport and seaport and the estimated population density across the distribution nodes.

The Tohoku earthquake, tsunami, and subsequent radiological incidents present a

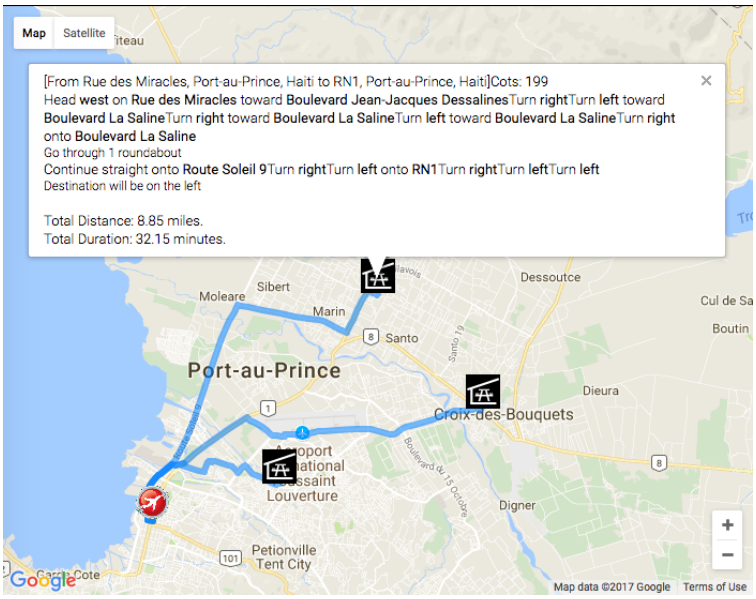


Figure 3.11: RealOpt-Regional: distribution paths and supply quantities.

glimpse of the devastating impacts of cascading failures and a catastrophic event in an advanced industrialized society. This is the first major hydrologic release of radiation isotopes and is the largest-ever release to the oceans. The cause was not a single event, but rather a cascading failure and the release persisted for a sustained period.

In Japan, while citizens have been educated about evacuation and response procedures for earthquakes and tsunamis, there is a serious knowledge gap regarding strategies and emergency guidelines for radiological emergencies, rapid screening and decontamination, even among emergency responders and citizens living close to the nuclear plants. During the Fukushima disaster, citizens learned very little through public media or from government officials. The mitigation efforts for radiation exposure and health risks were further clouded as a result of conflicting information released by the Japanese and the US government (e.g., on the radius for evacuation).

First month after March 11, 2011, over 3,000 workers were working around-the-clock on various emergency issues related to the failed nuclear plants. Lack of planning led to

missed screening of workers. To facilitate the disaster response effort, RealOpt-Regional was used to determine the sites and the numbers of shelters needed to house the displaced population, while RealOpt-CRC was deployed for rapid screening, decontamination and health registry and monitoring for emergency workers and the affected population.

RealOpt-CRC was used to design the optimal process layout for screening and decontamination. Some key tasks include detection of the presence of radioactive contamination on the body or clothing, removal of the contamination (decontamination), detection of the intake of radioactive materials into the body, registration/tracking, and medical consultation. The screening sites must include staff members and equipment capable of detecting contamination through beta/gamma portal monitors; monitoring for general contamination using hand-held instruments, fielding questions and addressing all concerns; and distributing event and follow-up information. Further, when internal contamination assessment is to be performed, licensed physicians are needed to carry out the procedure properly. Table 3.1 contrasts the results using two screening layouts, namely by switching the ordering on the use of portal monitor and hand-held screening instruments. The results underscore the importance of optimizing process design and the resource allocation, which has significant impact on the overall operations efficiency and systems performance.

Table 3.1: RealOpt-CRC: contrasting throughput for two different layouts for a screening and decontamination center.

	Layout 1	**Layout 2**
Throughput in 12 hours	52,982	40,217
Ave. time spent	8 minutes	15 minutes
Greeting	8 workers	6 workers
Paperwork	8 workers	6 workers
Portal monitor	16	3
Hand-held device	12	43
Decontamination	46	34
Bioassays	10 stations	8 stations

3.5 Challenges and Conclusions

Modeling and optimizing public health infrastructure involve elements of resource allocation under risk, uncertainty, and time pressure; large-scale supply-chain management; transportation and operational logistics; and medical treatment and population protection. The operations must be supported by an effective communication infrastructure. There is a necessity for vertical and horizontal integration and communication, where federal, state, local, tribal, territorial, private and business stakeholders work towards a common goal of a resilient public health system. The infrastructure must be flexible, scalable, and sustainable to support an effective and timely response, and to mount rapid recovery and mitigation operations.

In this study, we discuss a novel interactive web-based public health "digital" informatics decision support system, RealOpt, a suite of tools for real-time emergency preparedness planning, medical countermeasures dispensing, and public health surveillance. We also highlight its public health information sharing, reverse reporting, incidents tracking and monitoring, and supply/resupply coordination capabilities. All these elements are critical and essential for successful and effective response operations. We introduce herein RealOpt's usage in the Haiti and Japan Tohoku Disaster response efforts, the Ebola operations for W. Africa, and the Zika containment strategies across various countries in the world. The Tohoku earthquake, tsunami, and its subsequent radiological incidents accentuate the devastating impacts of cascading failures. One specific tragedy underscores the importance of real-time on-the-fly planning and optimization during an unfolding event: On March 11, 30 response officials gathered on the upper floor and roof in the disaster management headquarters for the town of Minamisanriku. The tsunami completely flooded the structure and only 11 people survived.

Coupled with Google Maps JavaScript API V3, RealOpt-Regional provides public health emergency preparedness coordinators spatial understanding of the planning region, empowers them in building a mathematical model for optimizing a service network in a graphical visualization environment; and, more importantly, it solves the problem in real time. The client-server architecture not only relieves the need for intensive computation from system users, it also embodies the concept of system modularization and encapsulation to maintain the extendibility of the solution engine without interfering with user experience.

Designing a safe and cost-effective network of service constructs (shelters, medical tents, point-of-dispensing PODs etc.) entails a complex process from planning to actual execution. RealOpt-Regional requires only a web browser for authorized users to access (including regional commanders, facility planners and managers, and epidemiologists), and provides a common platform for public health workers from different governmental levels (federal/state/local agencies). Using the RealOpt environment, stakeholders crucial to response planning and execution can share information, interact and integrate their respective components for decision/policy making within a trusted shared information environment.

RealOpt has been deployed to public health and emergency planners since December 2009. Currently there are over 200 regional commanders using RealOpt-Regional, most of the users are regional program directors or state response coordinators in Public Health or Emergency Management departments. There are over 10,000 RealOpt-POD site users. RealOpt has been used for response to hurricane, fire, flood, and other public health critical missions including international disaster crises. Its use helps in protecting and saving human lives.

We are currently integrating the standalone RealOpt-Evacuate into the online RealOpt

system. RealOpt-Evacuate is a large-scale evacuation system that serves as a strategic-based decision analysis support system in assisting regional evacuation operations, planning and training, in conjunction with population protection (medical countermeasures, medical care, shelters and food supplies). We also distribute a standalone RealOpt-RSS system, a software modeling, simulation and optimization tool for efficient logistics and operations management of Receipt, Stage and Storage (RSS) facilities and regional distribution nodes for medical countermeasures or other emergency supplies. RealOpt-RSS assists in planning quick and efficient strategies for processing medicines and/or equipment from the receiving point of the RSS to RSS vehicles for delivery. The software includes a warehouse design module, and specific computational modules related to the RSS facility, material handling equipment, staffing, and site configuration.

CHAPTER 4
CONCLUDING REMARKS

Infectious diseases remain a major threat to the public health nowadays and accurate prediction and immediate response at the initial stage of outbreaks are crucial to achieve early containment and minimize total infections. The topic of this dissertation is to use mathematical and computational methods to improve and advance the public health informatics, especially in terms of controlling infectious disease outbreaks. We build a general-purpose mathematical modeling framework for infectious disease outbreaks and use this modeling framework to assist decision-making to achieve rapid containment of disease outbreaks. We develop a real-time computerized decision support system for optimizing resource allocation and operations during emergency responses. This work starts with a theoretical mathematical model for epidemiology which is highly flexible and generalizable, then shows how this modeling framework can be used in practice to facilitate decision-making when combating with infectious disease outbreaks, and finally translates the theoretical model to practical decision support tools for first responders.

Mathematical models with system dynamics and interactions are powerful tools to analyze the trend of disease outbreaks and support real-time decision-making. However, building epidemiology models from scratch during the emergency is not efficient and may cause potential delays in response. We propose a general-purpose modeling framework for the spread of infectious diseases. This modeling framework extends the traditional compartmental models by categorizing each stage or compartment based on two different criteria: passive or active; vulnerable, contagious, or unaffected, and establishing relationships between stages based on their categorizations. We discuss models for two types of infectious diseases: contact-based and vector-borne diseases. The modeling technique for

vector-borne diseases can be generalized to depict the system dynamics for more than two species, given the well-defined inter-group disease contact structures. We derive the basic reproduction number of the system based on the method of the next-generation matrix. The spread of infectious disease has intrinsic stochasticity due to the randomness in the behavior of individuals and the heterogeneity among different population groups. Although the proposed modeling framework is deterministic, stochasticity can also be introduced to it using Langevin dynamics. All classical compartmental models can be derived from the proposed metamodel, which describes the dynamics of the system with a compact formulation and can be easily generalized and customized based on the biological properties and environmental features of different infectious disease outbreaks. The proposed modeling framework provides a high-level metamodel for epidemiology and enables the emergency responders to rapidly build and evaluate disease models and evaluate the performance of different containment strategies without the knowledge of the underlying mathematics.

We discussed how medical interventions can be introduced to the modeling framework of epidemiology both implicitly and explicitly. If the process of medical interventions can be treated as a black box, we can create new stages corresponding to the intervention process by assigning proper values in all model matrices. On the other hand, we can model the interventions in detail by segmenting the population into different groups. For example, to model the effect of vaccination, we can separate the population into two groups: outside and inside vaccine clinics. For the population outside the vaccine clinics, the previous disease propagation model will be used; for the population inside the vaccine clinics, besides the natural propagation of the disease, the vaccination effect will also drive the transition between disease stages. This modeling method is suitable to study the effect of the operations of healthcare facilities and point-of-dispensing sites (PODs) during emergency outbreaks and identify the optimal operation strategies. Based on the models for human interventions, optimization problems can be formulated to determine the optimal allocation

of medical resources and compare the effects of different containment strategies. Instead of treating all model matrices as constants, they can be modeled as time-variant variables and the values are controllable through human interventions at a certain cost. Then optimization problems can be formulated and solved to achieve containment under various constraints, assisting on-the-ground operations for disaster relief effort.

One major challenge in the application of this modeling framework is to accurately estimate model parameters. With the flexibility of the model, it can incorporate a very large number of parameters to describe the system dynamics in high resolutions. Some parameters can be obtained through the biological properties of the diseases, but other parameters require estimations. With a high degree of freedom, improper estimations of model parameters may result in over-fitting. In practice, decision-makers should start with simple but robust settings of the model and extend it to more complicated ones when such data are available. Our previous research shows that simple models have sufficient predicting accuracy in many cases and should be preferred when accurate parameter estimations are not immediately available. Sensitivity analysis is also useful to determine the relevant importance of different model parameters. One future research direction is to establish a data pipeline to populate the model parameters or use machine learning techniques to estimate model parameters. This can further automate the model fitting and evaluation process and improve the robustness and accuracy of the prediction.

To facilitate the decision-making process during public health emergencies in real time, we demonstrate the RealOpt suite, which consists of multiple modules to assist end-users at different levels in formulating and solving difficult mathematical problems including facility location and resource allocation. Public health workers and emergency responders without relative background knowledge can easily navigate through the graphical user interface of the system, build models that reflect their respective situations, and obtain solutions as

their guidance of designing and implementing emergency response plans. Additionally, with multiple RealOpt modules linked together, users can design the plan hierarchically from regional facility location optimization and logistics to the operational strategies and inventory management of each individual dispensing site. This allows end-users at multiple levels to work on the same emergency response plan under a unified framework and improves the overall efficiency.

Though the proposed modeling framework is highly generic and generalizable, there are limitations. For example, the current setup does not allow stages to have outflows caused by disease propagation and new infections combined. We sacrifice the capability of modeling these edge cases for the simplicity of our modeling framework, though this problem can be solved by defining two separate disease transition matrices. We point out that more complicated models can still be expressed under our framework with carefully quantified interactions between stages. Compartmental models and the corresponding ordinary differential equations are not difficult to solve numerically in general, but the real-time requirement of emergency response demands fast but accurate solutions. More theoretical and computational investigations on this ODE-based modeling framework would be necessary to further improve its performance and promote its usage. Approximation algorithm and interpolation method are also potential solutions to achieve fast and accurate analyses under the current framework. To cope with increasingly complicated scenarios in modern public health emergencies, the computerized tools used in this field also need to be updated continuously to reflect the recent development in emergency response and the complexity of the situations. Besides the improvements in theoretical models and algorithms, emergency response tools will also need to incorporate multiple data sources to better understand the current situation and predict future trends for infectious disease outbreaks. Crowd-sourced surveillance is becoming increasingly important with the advances of social networking. With proper technologies, these data can serve as key inputs to the RealOpt suite to provide

better model parameter estimations and more accurate intervention strategy evaluations.

REFERENCES

[1] H. Abbey, "An examination of the Reed-Frost theory of epidemics," *Human Biology*, vol. 24, no. 3, pp. 201–233, 1952.

[2] E. M. Abdelwhab and H. M. Hafez, "An overview of the epidemic of highly pathogenic H5N1 avian influenza virus in Egypt: Epidemiology and control challenges," *Epidemiology & Infection*, vol. 139, no. 5, pp. 647–657, 2011.

[3] E. Ahmed and A. S. Elgazzar, "On fractional order differential equations model for nonlocal epidemics," *Physica A: Statistical Mechanics and its Applications*, vol. 379, no. 2, pp. 607–614, 2007.

[4] O. E. Aiello and M. A. A. da Silva, "New approach to dynamical Monte Carlo methods: Application to an epidemic model," *Physica A: Statistical Mechanics and its Applications*, vol. 327, no. 3-4, pp. 525–534, 2003.

[5] D. J. Alexander, "A review of avian influenza in different bird species," *Veterinary Microbiology*, vol. 74, no. 1–2, pp. 3–13, 2000.

[6] K. Alibek, "Smallpox: A disease and a weapon," *International Journal of Infectious Diseases*, vol. 8S2, pp. S3–S8, 2004.

[7] E. Allen, *Modeling with Itô Stochastic Differential Equations*. Dordrecht: Springer, 2007.

[8] L. J. S. Allen, *An Introduction to Mathematical Biology*. Upper Saddle River: Pearson, 2007.

[9] L. J. S. Allen and A. M. Burgin, "Comparison of deterministic and stochastic SIS and SIR models in discrete time," *Mathematical Biosciences*, vol. 163, no. 1, pp. 1–33, 2000.

[10] C. L. Althaus, "Estimating the reproduction number of Ebola virus (EBOV) during the 2014 outbreak in West Africa," *PLoS Currents Outbreaks*, Sep 2, 2014.

[11] R. M. Anderson and R. M. May, Eds., *Population Biology of Infectious Diseases, Report of the Dahlem Workshop on Population Biology of Infectious Disease Agents, Berlin 1982, March 14-19*, Berlin, Heidelberg: Springer, 1982.

[12] H. Andersson and T. Britton, *Stochastic Epidemic Models and Their Statistical Analysis*. New York: Springer-Verlag, 2000.

[13] J. Arino, F. Brauer, P. Van den Driessche, J. Watmough, and J. Wu, "A model for influenza with vaccination and antiviral treatment," *Journal of Theoretical Biology*, vol. 253, no. 1, pp. 118–130, 2008.

[14] J. Arino, R. Jordan, and P. Van den Driessche, "Quarantine in a multi-species epidemic model with spatial dynamics," *Mathematical Biosciences*, vol. 206, no. 1, pp. 46–60, 2007.

[15] A. W. Artenstein, C. Johnson, T. C. Marbury, D. Morrison, P. S. Blum, *et al.*, "A novel, cell culture-derived smallpox vaccine in vaccinia-naive adults," *Vaccine*, vol. 23, no. 25, pp. 3301–3309, 2005.

[16] N. T. J. Bailey, *The Mathematical Theory of Infectious Diseases and Its Applications*, 2nd ed. London: Griffin, 1975.

[17] S. Baize, D. Pannetier, L. Oestereich, T. Rieger, L. Koivogui, *et al.*, "Emergence of Zaire Ebola virus disease in Guinea," *New England Journal of Medicine*, vol. 371, no. 15, pp. 1418–1425, 2014.

[18] D. Balcan, B. Gonçalves, H. Hu, J. J. Ramasco, V. Colizza, *et al.*, "Modeling the spatial spread of infectious diseases: The global epidemic and mobility computational model," *Journal of Computational Science*, vol. 1, no. 3, pp. 132–145, 2010.

[19] F. Ball and P. O'Neill, "The distribution of general final state random variables for stochastic epidemic models," *Journal of Applied Probability*, vol. 36, no. 2, pp. 473–491, 1999.

[20] S. Bansal, B. T. Grenfell, and L. A. Meyers, "When individual behaviour matters: Homogeneous and network models in epidemiology," *Journal of the Royal Society Interface*, vol. 4, no. 16, pp. 879–891, 2007.

[21] I. Bargielowski, D. Nimmo, L. Alphey, and J. C. Koella, "Comparison of life history characteristics of the genetically modified OX513A line and a wild type strain of Aedes aegypti," *PLoS One*, vol. 6, no. 6, e20699, 2011.

[22] M. S. Bartlett, "Deterministic and stochastic models for recurrent epidemics," in *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Volume 4: Contributions to Biology and Problems of Health*, J. Neyman, Ed., Berkeley: University of California Press, 1956, pp. 81–109.

[23] R. H. Beach, C. Poulos, and S. K. Pattanayak, "Farm economics of bird flu," *Canadian Journal of Agricultural Economics*, vol. 55, no. 4, pp. 471–483, 2007.

[24] W. G. C. Bearcroft, "Zika virus infection experimentally induced in a human volunteer," *Transactions of the Royal Society of Tropical Medicine and Hygiene*, vol. 50, no. 5, pp. 442–448, 1956.

[25] W. B. Becker, "The isolation and classification of Tern virus: Influenza virus A/Tern/South Africa/1961," *The Journal of Hygiene*, vol. 64, no. 3, pp. 309–320, 1966.

[26] J. H. Beigel, J. Farrar, A. M. Han, F. G. Hayden, R. Hyer, *et al.*, "Avian influenza A (H5N1) infection in humans," *The New England Journal of Medicine*, vol. 353, no. 13, pp. 1374–1385, 2005.

[27] R. J. Bellamy and A. R. Freedman, "Bioterrorism," *QJM: An International Journal of Medicine*, vol. 94, no. 4, pp. 227–234, 2001.

[28] D. Bernoulli and S. Blower, "An attempt at a new analysis of the mortality caused by smallpox and of the advantages of inoculation to prevent it," *Reviews in Medical Virology*, vol. 14, no. 5, pp. 275–288, 2004.

[29] S. Binder, A. M. Levitt, J. J. Sacks, and J. M. Hughes, "Emerging infectious diseases: Public health issues for the 21st century," *Science*, vol. 284, no. 5418, pp. 1311–1313, 1999.

[30] I. I. Bogoch, O. J. Brady, M. U. Kraemer, M. German, M. I. Creatore, *et al.*, "Anticipating the international spread of Zika virus from Brazil," *The Lancet*, vol. 387, no. 10016, pp. 335–336, 2016.

[31] S. A. Bozzette, R. Boer, V. Bhatnagar, J. L. Brower, E. B. Keeler, *et al.*, "A model for a smallpox-vaccination policy," *New England Journal of Medicine*, vol. 348, no. 5, pp. 416–425, 2003.

[32] O. Brady, S. Hay, and P. Horby, "Scale up the supply of experimental Ebola drugs," *Nature*, vol. 512, no. 7514, p. 233, 2014.

[33] P. Brasil, G. A. Calvet, A. M. Siqueira, M. Wakimoto, P. C. de Sequeira, *et al.*, "Zika virus outbreak in Rio de Janeiro, Brazil: Clinical characterization, epidemiological and virological aspects," *PLoS Neglected Tropical Diseases*, vol. 10, no. 4, e0004636, 2016.

[34] F. Brauer, "Mathematical epidemiology is not an oxymoron," *BMC Public Health*, vol. 9S1, p. S2, 2009.

[35] ——, "Models for the spread of universally fatal diseases," *Journal of Mathematical Biology*, vol. 28, no. 4, pp. 451–462, 1990.

[36] D. Breda, O. Diekmann, W. F. de Graaf, A. Pugliese, and R. Vermiglio, "On the formulation of epidemic models (an appraisal of Kermack and McKendrick)," *Journal of Biological Dynamics*, vol. 6, no. S2, pp. 103–117, 2012.

[37] T. Britton, "Stochastic epidemic models: A survey," *Mathematical Biosciences*, vol. 225, no. 1, pp. 24–35, 2010.

[38] E. Camacho, M. Paternina-Gomez, P. J. Blanco, J. E. Osorio, and M. T. Aliota, "Detection of autochthonous Zika virus transmission in Sincelejo, Colombia," *Emerging Infectious Diseases*, vol. 22, no. 5, pp. 927–929, 2016.

[39] G. S. Campos, A. C. Bandeira, and S. I. Sardi, "Zika virus outbreak, Bahia, Brazil," *Emerging Infectious Diseases*, vol. 21, no. 10, pp. 1885–1886, 2015.

[40] V.-M. Cao-Lormeau, C. Roche, A. Teissier, E. Robin, A.-L. Berry, *et al.*, "Zika virus, French Polynesia, South Pacific, 2013," *Emerging Infectious Diseases*, vol. 20, no. 6, pp. 1085–1086, 2014.

[41] V. Capasso and G. Serio, "A generalization of the Kermack-McKendrick deterministic epidemic model," *Mathematical Biosciences*, vol. 42, no. 1–2, pp. 43–61, 1978.

[42] E. P. Carlin, N. Giller, and R. Katz, "Estimating the size of the US population at risk of severe adverse events from replicating smallpox vaccine," *Public Health Nursing*, vol. 34, no. 3, pp. 200–209, 2017.

[43] F. Carrat, J. Luong, H. Lao, A.-V. Sallé, C. Lajaunie, *et al.*, "A 'small-world-like' model for comparing interventions aimed at preventing and controlling influenza pandemics," *BMC Medicine*, vol. 4, p. 26, 2006.

[44] D. O. Carvalho, A. R. McKemey, L. Garziera, R. Lacroix, C. A. Donnelly, *et al.*, "Suppression of a field population of Aedes aegypti in Brazil by sustained release of transgenic male mosquitoes," *PLoS Neglected Tropical Diseases*, vol. 9, no. 7, e0003864, 2015.

[45] C. Castillo-Chavez, Z. Feng, and W. Huang, "On the computation of R0 and its role on global stability," in *Mathematical Approaches for Emerging and Reemerging Infectious Diseases: An Introduction*, C. Castillo-Chavez, S. Blower, P. Van den Driessche, D. Kirschner, and A.-A. Yakubu, Eds., New York: Springer-Verlag, 2002, pp. 229–250.

[46] Centers for Disease Control and Prevention. (2017). Zika virus, [Online]. Available: `https://www.cdc.gov/zika/` (visited on 01/31/2018).

[47] M. Chan, "Ebola virus disease in West Africa — No early end to the outbreak," *New England Journal of Medicine*, vol. 371, no. 13, pp. 1183–1185, 2014.

[48] G. Chen and T. Li, "Stability of stochastic delayed SIR model," *Stochastics and Dynamics*, vol. 9, no. 2, pp. 231–252, 2009.

[49] W.-Y. Chen and S. Bokka, "Stochastic modeling of nonlinear epidemiology," *Journal of Theoretical Biology*, vol. 234, no. 4, pp. 455–470, 2005.

[50] G. Chowell, N. W. Hengartner, C. Castillo-Chavez, P. W. Fenimore, and J. M. Hyman, "The basic reproductive number of Ebola and the effects of public health measures: The cases of Congo and Uganda," *Journal of Theoretical Biology*, vol. 229, no. 1, pp. 119–126, 2004.

[51] D. Clancy and A. B. Piunovskiy, "An explicit optimal isolation policy for a deterministic epidemic model," *Applied Mathematics and Computation*, vol. 163, no. 3, pp. 1109–1121, 2005.

[52] V. Colizza, A. Barrat, M. Barthélemy, A.-J. Valleron, and A. Vespignani, "Modeling the worldwide spread of pandemic influenza: Baseline case and containment interventions," *PLoS Medicine*, vol. 4, no. 1, e13, 2007.

[53] V. Colizza, A. Barrat, M. Barthélemy, and A. Vespignani, "The modeling of global epidemics: Stochastic dynamics and predictability," *Bulletin of Mathematical Biology*, vol. 68, no. 8, pp. 1893–1921, 2006.

[54] J. M. Conway, A. R. Tuite, D. N. Fisman, N. Hupert, R. Meza, *et al.*, "Vaccination against 2009 pandemic H1N1 in a population dynamical model of Vancouver, Canada: Timing is everything," *BMC Public Health*, vol. 11, p. 932, 2011.

[55] B. Cooper and M. Lipsitch, "The analysis of hospital infection data using hidden Markov models," *Biostatistics*, vol. 5, no. 2, pp. 223–237, 2004.

[56] S. E. Cosgrove, N. O. Fishman, T. R. Talbot, K. F. Woeltje, W. Schaffner, *et al.*, "Strategies for use of a limited influenza vaccine supply," *The Journal of the American Medical Association*, vol. 293, no. 2, pp. 229–232, 2005.

[57] A. Costero, J. D. Edman, G. G. Clark, and T. W. Scott, "Life table study of Aedes aegypti (Diptera: Culicidae) in Puerto Rico fed only human blood versus blood plus sugar," *Journal of Medical Entomology*, vol. 35, no. 5, pp. 809–813, 1998.

[58] N. J. Cox, S. E. Tamblyn, and T. Tam, "Influenza pandemic planning," *Vaccine*, vol. 21, no. 16, pp. 1801–1803, 2003.

[59] N. Dalal, D. Greenhalgh, and X. Mao, "A stochastic model of AIDS and condom use," *Journal of Mathematical Analysis and Applications*, vol. 325, no. 1, pp. 36–53, 2007.

[60] D. J. Daley and J. Gani, *Epidemic Modelling: An Introduction*. Cambridge: Cambridge University Press, 1999.

[61] M. Derouich and A. Boutayeb, "Dengue fever: Mathematical modelling and computer simulation," *Applied Mathematics and Computation*, vol. 177, no. 2, pp. 528–544, 2006.

[62] G. W. A. Dick, S. F. Kitchen, and A. J. Haddow, "Zika virus (I). Isolations and serological specificity," *Transactions of the Royal Society of Tropical Medicine and Hygiene*, vol. 46, no. 5, pp. 509–520, 1952.

[63] O. Diekmann and J. A. P. Heesterbeek, *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation*. New York: John Wiley & Sons, 2000.

[64] O. Diekmann, J. A. P. Heesterbeek, and J. A. J. Metz, "On the definition and the computation of the basic reproduction ratio R0 in models for infectious diseases in heterogeneous populations," *Journal of Mathematical Biology*, vol. 28, no. 4, pp. 365–382, 1990.

[65] O. Diekmann, M. C. M. de Jong, and J. A. J. Metz, "A deterministic epidemic model taking account of repeated contacts between the same individuals," *Journal of Applied Probability*, vol. 35, no. 2, pp. 448–462, 1998.

[66] K. Dietz, "The first epidemic model: A historical note on P. D. En'ko," *Australian & New Zealand Journal of Statistics*, vol. 30A, no. 1, pp. 56–65, 1988.

[67] Y. Ding, M. Xu, and L. Hu, "Asymptotic behavior and stability of a stochastic model for AIDS transmission," *Applied Mathematics and Computation*, vol. 204, no. 1, pp. 99–108, 2008.

[68] E. Dion, L. VanSchalkwyk, and E. F. Lambin, "The landscape epidemiology of foot-and-mouth disease in South Africa: A spatially explicit multi-agent simulation," *Ecological Modelling*, vol. 222, no. 13, pp. 2059–2072, 2011.

[69] D. S. Disraelly, T. J. Walsh, and C. A. Curling, "A new methodology for estimating contagious biological agent casualties as a function of time," *Mathematical and Computer Modelling*, vol. 54, no. 1–2, pp. 649–658, 2011.

[70] J. M. Drake, R. B. Kaul, L. W. Alexander, S. M. O'Regan, A. M. Kramer, *et al.*, "Ebola cases and health system demand in Liberia," *PLoS Biology*, vol. 13, no. 1, e1002056, 2015.

[71] P. Van den Driessche and J. Watmough, "Further notes on the basic reproduction number," in *Mathematical Epidemiology. Lecture Notes in Mathematics*, F. Brauer, P. Van den Driessche, and J. Wu, Eds., vol. 1945, Berlin, Heidelberg: Springer, 2008, pp. 159–178.

[72] ——, "Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission," *Mathematical Biosciences*, vol. 180, no. 1-2, pp. 29–48, 2002.

[73] M. R. Duffy, T.-H. Chen, W. T. Hancock, A. M. Powers, J. L. Kool, *et al.*, "Zika virus outbreak on Yap Island, Federated States of Micronesia," *New England Journal of Medicine*, vol. 360, no. 24, pp. 2536–2543, 2009.

[74] EFSA Panel on Animal Health and Welfare, "Scientific opinion on the influence of genetic parameters on the welfare and the resistance to stress of commercial broilers," *EFSA Journal*, vol. 8, no. 7, p. 1666, 2010.

[75] M. Eichner and K. Dietz, "Transmission potential of smallpox: Estimates based on detailed data from an outbreak," *American Journal of Epidemiology*, vol. 158, no. 2, pp. 110–117, 2003.

[76] M. Eichner, S. F. Dowell, and N. Firese, "Incubation period of Ebola hemorrhagic virus subtype Zaire," *Osong Public Health and Research Perspectives*, vol. 2, no. 1, pp. 3–7, 2011.

[77] E. J. Emanuel and A. Wertheimer, "Who should get influenza vaccine when not all can?" *Science*, vol. 312, no. 5775, pp. 854–855, 2006.

[78] R. A. Erickson, "Impact of climate change on the population dynamics of Aedes albopictus and disease dynamics of dengue," Master's thesis, Texas Tech University, Lubbock, 2009.

[79] R. A. Erickson, S. M. Presley, L. J. S. Allen, K. R. Long, and S. B. Cox, "A dengue model with a dynamic Aedes albopictus vector population," *Ecological Modelling*, vol. 221, no. 24, pp. 2899–2908, 2010.

[80] L. Esteva and M. Matias, "A model for vector transmitted diseases with saturation incidence," *Journal of Biological Systems*, vol. 9, no. 4, pp. 235–245, 2001.

[81] L. Esteva and C. Vargas, "A model for dengue disease with variable human population," *Journal of Mathematical Biology*, vol. 38, no. 3, pp. 220–240, 1999.

[82] ——, "Analysis of a dengue disease transmission model," *Mathematical Biosciences*, vol. 150, no. 2, pp. 131–151, 1998.

[83] M. J. Faddy, "Nonlinear stochastic compartmental models," *Mathematical Medicine and Biology*, vol. 2, no. 4, pp. 287–297, 1985.

[84] Z. Feng and J. X. Velasco-Hernández, "Competitive exclusion in a vector-host model for the dengue fever," *Journal of Mathematical Biology*, vol. 35, no. 5, pp. 523–544, 1997.

[85] F. Fenner, D. A. Henderson, I. Arita, Z. Jezek, and I. D. Ladnyi, *Smallpox and Its Eradication*. Geneva, Switzerland: World Health Organization, 1988.

[86] N. M. Ferguson, C. Fraser, C. A. Donnelly, A. C. Ghani, and R. M. Anderson, "Public health risk from the avian H5N1 influenza epidemic," *Science*, vol. 304, no. 5673, pp. 968–969, 2004.

[87] N. M. Ferguson, M. J. Keeling, W. J. Edmunds, R. Gani, B. T. Grenfell, *et al.*, "Planning for smallpox outbreaks," *Nature*, vol. 425, no. 6959, pp. 681–685, 2003.

[88] D. N. Fisman, E. Khoo, and A. R. Tuite, "Early epidemic dynamics of the West African 2014 Ebola outbreak: Estimates derived with a simple two-parameter model," *PLoS Currents Outbreaks*, Sep 8, 2014.

[89] D. A. Focks, E. Daniels, D. G. Haile, and J. E. Keesling, "A simulation model of the epidemiology of urban dengue fever: Literature analysis, model development, preliminary validation, and samples of simulation results," *The American Journal of Tropical Medicine and Hygiene*, vol. 53, no. 5, pp. 489–506, 1995.

[90] B. D. Foy, K. C. Kobylinski, J. L. C. Foy, B. J. Blitvich, A. T. da Rosa, *et al.*, "Probable non-vector-borne transmission of Zika virus, Colorado, USA," *Emerging Infectious Diseases*, vol. 17, no. 5, pp. 880–882, 2011.

[91] S. E. Frey, F. K. Newman, J. Cruz, W. B. Shelton, J. M. Tennant, *et al.*, "Dose-related effects of smallpox vaccine," *New England Journal of Medicine*, vol. 346, no. 17, pp. 1275–1280, 2002.

[92] E. Fukuda, S. Kokubo, J. Tanimoto, Z. Wang, A. Hagishima, *et al.*, "Risk assessment for infectious disease and its impact on voluntary vaccination behavior in social networks," *Chaos, Solitons & Fractals*, vol. 68, pp. 1–9, 2014.

[93] T. Garske, P. Clarke, and A. C. Ghani, "The transmissibility of highly pathogenic avian influenza in commercial poultry in industrialised countries," *PLoS One*, vol. 2, no. 4, e349, 2007.

[94] M. T. Gastner and M. E. J. Newman, "Diffusion-based method for producing density-equalizing maps," *Proceedings of the National Academy of Sciences*, vol. 101, no. 20, pp. 7499–7504, 2004.

[95] M. D. Gershman, K. M. Angelo, J. Ritchey, D. P. Greenberg, R. D. Muhammad, *et al.*, "Addressing a yellow fever vaccine shortage — United States, 2016–2017," *Morbidity and Mortality Weekly Report*, vol. 66, no. 17, pp. 457–459, 2017.

[96] G. J. Gibson and E. Renshaw, "Estimating parameters in stochastic compartmental models using Markov chain methods," *Mathematical Medicine and Biology*, vol. 15, no. 1, pp. 19–40, 1998.

[97] D. T. Gillespie, "The chemical Langevin equation," *The Journal of Chemical Physics*, vol. 113, no. 1, pp. 297–306, 2000.

[98] M. F. C. Gomes, A. P. y Piontti, L. Rossi, D. Chao, I. M. Longini Jr, *et al.*, "Assessing the international spreading risk associated with the 2014 West African Ebola outbreak," *PLoS Currents Outbreaks*, Sep 2, 2014.

[99] Government Accountability Office, "National capital region — 2010 strategic plan is generally consistent with characteristics of effective strategies," United States Government Accountability Office, Washington, D.C., Tech. Rep., 2011.

[100] G. Grard, M. Caron, I. M. Mombo, D. Nkoghe, S. M. Ondo, *et al.*, "Zika virus in Gabon (Central Africa) — 2007: A new threat from Aedes albopictus?" *PLoS Neglected Tropical Diseases*, vol. 8, no. 2, e2681, 2014.

[101] N. C. Grassly and C. Fraser, "Mathematical models of infectious disease transmission," *Nature Reviews Microbiology*, vol. 6, pp. 477–487, 2008.

[102] A. Gray, D. Greenhalgh, L. Hu, X. Mao, and J. Pan, "A stochastic differential equation SIS epidemic model," *SIAM Journal on Applied Mathematics*, vol. 71, no. 3, pp. 876–902, 2011.

[103] M. Greenwood, "On the statistical measure of infectiousness," *The Journal of Hygiene*, vol. 31, no. 3, pp. 336–351, 1931.

[104] V. J. Haas, A. Caliri, and M. A. A. da Silva, "Temporal duration and event size distribution at the epidemic threshold," *Journal of Biological Physics*, vol. 25, no. 4, pp. 309–324, 1999.

[105] A. D. Haddow, A. J. Schuh, C. Y. Yasuda, M. R. Kasper, V. Heang, *et al.*, "Genetic characterization of Zika virus strains: Geographic expansion of the Asian lineage," *PLoS Neglected Tropical Diseases*, vol. 6, no. 2, e1477, 2012.

[106] M. E. Halloran, I. M. Longini Jr, A. Nizam, and Y. Yang, "Containing bioterrorist smallpox," *Science*, vol. 298, no. 5597, pp. 1428–1432, 2002.

[107] W. H. Hamer, "Epidemic disease in England: The evidence of variability and of persistency of type," *The Lancet*, vol. 167, no. 4306, pp. 655–662, 1906.

[108] E. B. Hayes, "Zika virus outside Africa," *Emerging Infectious Diseases*, vol. 15, no. 9, pp. 1347–1350, 2009.

[109] D. A. Henderson, "Smallpox: Clinical and epidemiologic features," *Emerging Infectious Diseases*, vol. 5, no. 4, pp. 537–539, 1999.

[110] H. W. Hethcote, "A thousand and one epidemic models," in *Frontiers in Mathematical Biology. Lecture Notes in Biomathematics*, S. A. Levin, Ed., vol. 100, Berlin, Heidelberg: Springer, 1994, pp. 504–515.

[111] H. Inaba, "On a pandemic threshold theorem of the early Kermack-McKendrick model with individual heterogeneity," *Mathematical Population Studies*, vol. 21, no. 2, pp. 95–111, 2014.

[112] S. Ioos, H.-P. Mallet, I. L. Goffart, V. Gauthier, T. Cardoso, *et al.*, "Current Zika virus epidemiology and recent epidemics," *Médecine et Maladies Infectieuses*, vol. 44, no. 7, pp. 302–307, 2014.

[113] V. Isham, "Stochastic models for epidemics," in *Celebrating Statistics: Papers in Honor of Sir David Cox on His 80th Birthday*, A. C. Davison, Y. Dodge, and N. Wermuth, Eds., Oxford: Oxford University Press, 2005, pp. 27–54.

[114] S. Iwami, Y. Takeuchi, and X. Liu, "Avian flu pandemic: Can we prevent it?" *Journal of Theoretical Biology*, vol. 257, no. 1, pp. 181–190, 2009.

[115] ——, "Avian-human influenza epidemic model," *Mathematical Biosciences*, vol. 207, no. 1, pp. 1–25, 2007.

[116] J. A. Jacquez and C. P. Simon, "The stochastic SI model with recruitment and deaths I. Comparison with the closed SIS model," *Mathematical Biosciences*, vol. 117, no. 1-2, pp. 77–125, 1993.

[117] M. Johnson and R. Patterson. (2017). Infectious disease collides with changing climate, [Online]. Available: `https://projects.jsonline.com/news/2017/5/24/infectious-disease-collides-with-changing-climate.html` (visited on 01/31/2018).

[118] E. H. Kaplan, D. L. Craft, and L. M. Wein, "Emergency response to a smallpox attack: The case for mass vaccination," *Proceedings of the National Academy of Sciences*, vol. 99, no. 16, pp. 10 935–10 940, 2002.

[119] S. Karsten, G. Rave, and J. Krieter, "Monte Carlo simulation of classical swine fever epidemics and control: I. General concepts and description of the model," *Veterinary Microbiology*, vol. 108, no. 3-4, pp. 187–198, 2005.

[120] J. R. Kaufmann, R. Miller, and J. Cheyne, "Vaccine supply chains need to be better funded and strengthened, or lives will be at risk," *Health Affairs*, vol. 30, no. 6, pp. 1113–1121, 2011.

[121] M. J. Keeling and K. T. D. Eames, "Networks and epidemic models," *Journal of the Royal Society Interface*, vol. 2, no. 4, pp. 295–307, 2005.

[122] D. G. Kendall, "Deterministic and stochastic epidemics in closed populations," in *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Volume 4: Contributions to Biology and Problems of Health*, J. Neyman, Ed., Berkeley: University of California Press, 1956, pp. 149–165.

[123] W. O. Kermack and A. G. McKendrick, "A contribution to the mathematical theory of epidemics — I. 1927," *Bulletin of Mathematical Biology*, vol. 53, no. 1-2, pp. 33–55, 1991.

[124] ——, "A contribution to the mathematical theory of epidemics — II. The problem of endemicity. 1932," *Bulletin of Mathematical Biology*, vol. 53, no. 1-2, pp. 57–87, 1991.

[125] ——, "A contribution to the mathematical theory of epidemics — III. Further studies of the problem of endemicity. 1933," *Bulletin of Mathematical Biology*, vol. 53, no. 1-2, pp. 89–118, 1991.

[126] K. I. Kim, Z. Lin, and L. Zhang, "Avian-human influenza epidemic model with diffusion," *Nonlinear Analysis: Real World Applications*, vol. 11, no. 1, pp. 313–322, 2010.

[127] M. A. Kiskowski, "A three-scale network model for the early growth dynamics of 2014 West Africa Ebola epidemic," *PLoS Currents Outbreaks*, Nov 13, 2014.

[128] Z. Kou, Y. Li, Z. Yin, S. Guo, M. Wang, *et al.*, "The survey of H5N1 flu virus in wild birds in 14 provinces of China from 2004 to 2007," *PLoS One*, vol. 4, no. 9, e6926, 2009.

[129] M. U. G. Kraemer, N. R. Faria, R. C. Reiner Jr, N. Golding, B. Nikolay, *et al.*, "Spread of yellow fever virus outbreak in Angola and the Democratic Republic of

the Congo 2015–16: A modelling study," *The Lancet Infectious Diseases*, vol. 17, no. 3, pp. 330–338, 2017.

[130] G. Kuno, G.-J. J. Chang, K. R. Tsuchiya, N. Karabatsos, and C. B. Cropp, "Phylogeny of the genus Flavivirus," *Journal of Virology*, vol. 72, no. 1, pp. 73–83, 1998.

[131] M. Kuperman and G. Abramson, "Small world effect in an epidemiological model," *Physical Review Letters*, vol. 86, no. 13, pp. 2909–2912, 2001.

[132] J. R. La Montagne and A. S. Fauci, "Intradermal influenza vaccination — Can less be more?" *New England Journal of Medicine*, vol. 351, no. 22, pp. 2330–2332, 2004.

[133] B. Y. Lee, S. T. Brown, P. C. Cooley, R. K. Zimmerman, W. D. Wheaton, *et al.*, "A computer simulation of employee vaccination to mitigate an influenza epidemic," *American Journal of Preventive Medicine*, vol. 38, no. 3, pp. 247–257, 2010.

[134] B. Y. Lee, S. T. Brown, G. W. Korch, P. C. Cooley, R. K. Zimmerman, *et al.*, "A computer simulation of vaccine prioritization, allocation, and rationing during the 2009 H1N1 influenza pandemic," *Vaccine*, vol. 28, no. 31, pp. 4875–4879, 2010.

[135] E. K. Lee, C.-H. Chen, F. H. Pietz, and B. A. Benecke, "Disease propagation analysis and mitigation strategies for effective mass dispensing," *AMIA Annual Symposium Proceedings*, vol. 2010, pp. 427–431, 2010.

[136] ——, "Modeling and optimizing the public-health infrastructure for emergency response," *Interfaces*, vol. 39, no. 5, pp. 476–490, 2009.

[137] E. K. Lee, Y. Liu, and F. H. Pietz, "A compartmental model for Zika virus with dynamic human and vector populations," *AMIA Annual Symposium Proceedings*, vol. 2016, pp. 743–752, 2016.

[138] E. K. Lee, S. Maheshwary, J. Mason, and W. Glisson, "Decision support system for mass dispensing of medications for infectious disease outbreaks and bioterrorist attacks," *Annals of Operations Research*, vol. 148, no. 1, pp. 25–53, 2006.

[139] ——, "Large-scale dispensing for emergency response to bioterrorism and infectious-disease outbreak," *Interfaces*, vol. 36, no. 6, pp. 591–607, 2006.

[140] E. K. Lee, H. I. Nakaya, F. Yuan, T. D. Querec, G. Burel, *et al.*, "Machine learning for predicting vaccine immunogenicity," *Interfaces*, vol. 46, no. 5, pp. 368–390, 2016.

[141] E. K. Lee, F. H. Pietz, B. A. Benecke, J. Mason, and G. Burel, "Advancing public health and medical preparedness with operations research," *Interfaces*, vol. 43, no. 1, pp. 79–98, 2013.

[142] E. K. Lee, H. K. Smalley, Y. Zhang, F. H. Pietz, and B. A. Benecke, "Facility location and multi-modality mass dispensing strategies and emergency response for biodefence and infectious disease outbreaks," *International Journal of Risk Assessment and Management*, vol. 12, no. 2–4, pp. 311–351, 2009.

[143] E. K. Lee, F. Yuan, F. H. Pietz, B. A. Benecke, and G. Burel, "Vaccine prioritization for effective pandemic response," *Interfaces*, vol. 45, no. 5, pp. 425–443, 2015.

[144] E. K. Lee and M. Zaider, "Operations research advances cancer therapeutics," *Interfaces*, vol. 38, no. 1, pp. 5–25, 2008.

[145] J. Legrand, R. F. Grais, P.-Y. Boelle, A.-J. Valleron, and A. Flahault, "Understanding the dynamics of Ebola epidemics," *Epidemiology & Infection*, vol. 135, no. 4, pp. 610–621, 2007.

[146] D. S. Lemons and A. Gythiel, "Paul Langevin's 1908 paper "On the theory of Brownian Motion"["Sur la théorie du mouvement brownien," C. R. Acad. Sci. (Paris) 146, 530–533 (1908)]," *American Journal of Physics*, vol. 65, no. 11, pp. 1079–1081, 1997.

[147] J. A. Lewnard, M. L. N. Mbah, J. A. Alfaro-Murillo, F. L. Altice, L. Bawo, *et al.*, "Dynamics and control of Ebola virus transmission in Montserrado, Liberia: A mathematical modelling analysis," *The Lancet Infectious Diseases*, vol. 14, no. 12, pp. 1189–1195, 2014.

[148] J. Li and Z. Ma, "Qualitative analyses of SIS epidemic model with vaccination and varying total population size," *Mathematical and Computer Modelling*, vol. 35, no. 11-12, pp. 1235–1243, 2002.

[149] J. Li and X. Zou, "Generalization of the Kermack-McKendrick SIR model to a patchy environment for a disease with latency," *Mathematical Modelling of Natural Phenomena*, vol. 4, no. 2, pp. 92–118, 2009.

[150] C.-M. Liao, C.-F. Chang, and H.-M. Liang, "A probabilistic transmission dynamic model to assess indoor airborne infection risks," *Risk Analysis*, vol. 25, no. 5, pp. 1097–1107, 2005.

[151] N. P. Lindsey, B. A. Schroeder, E. R. Miller, M. M. Braun, A. F. Hinckley, *et al.*, "Adverse event reports following yellow fever vaccination," *Vaccine*, vol. 26, no. 48, pp. 6077–6082, 2008.

[152] X. Liu, Y. Takeuchi, and S. Iwami, "SVIR epidemic models with vaccination strategies," *Journal of Theoretical Biology*, vol. 253, no. 1, pp. 1–11, 2008.

[153] Q. Lu, "Stability of SIRS system with random perturbations," *Physica A: Statistical Mechanics and its Applications*, vol. 388, no. 18, pp. 3677–3686, 2009.

[154] M. O. Lwin, S. Vijaykumar, O. N. N. Fernando, S. A. Cheong, V. S. Rathnayake, *et al.*, "A 21st century approach to tackling dengue: Crowdsourced surveillance, predictive mapping and tailored communication," *Acta Tropica*, vol. 130, pp. 100–107, 2014.

[155] M. S. Massoudi, L. Barker, and B. Schwartz, "Effectiveness of postexposure vaccination for the prevention of smallpox: Results of a delphi analysis," *The Journal of Infectious Diseases*, vol. 188, no. 7, pp. 973–976, 2003.

[156] R. M. May and A. L. Lloyd, "Infection dynamics on scale-free networks," *Physical Review E*, vol. 64, no. 6, p. 066112, 2001.

[157] B. McCloskey, O. Dar, A. Zumla, and D. L. Heymann, "Emerging infectious diseases and pandemic potential: Status quo and reducing risk of global spread," *The Lancet Infectious Diseases*, vol. 14, no. 10, pp. 1001–1010, 2014.

[158] A. G. McKendrick, "Applications of mathematics to medical problems," *Proceedings of the Edinburgh Mathematical Society*, vol. 44, pp. 98–130, 1925.

[159] J. Medlock and A. P. Galvani, "Optimizing influenza vaccine distribution," *Science*, vol. 325, no. 5948, pp. 1705–1708, 2009.

[160] J. Medlock and L. A. Meyers, "Optimizing allocation for a delayed influenza vaccination campaign," *PLoS Currents Influenza*, Dec 13, 2009.

[161] M. I. Meltzer, N. J. Cox, and K. Fukuda, "The economic impact of pandemic influenza in the United States: Priorities for intervention," *Emerging Infectious Diseases*, vol. 5, no. 5, pp. 659–671, 1999.

[162] M. A. Miller, C. Viboud, D. R. Olson, R. F. Grais, M. A. Rabaa, *et al.*, "Prioritization of influenza pandemic vaccination to minimize years of life lost," *The Journal of Infectious Diseases*, vol. 198, no. 3, pp. 305–311, 2008.

[163] G. J. Milne, J. K. Kelso, H. A. Kelly, S. T. Huband, and J. McVernon, "A small community model for the transmission of infectious diseases: Comparison of school closure as an intervention in individual-based models of an influenza pandemic," *PloS One*, vol. 3, no. 12, e4005, 2008.

[164] D. M. Morens and A. S. Fauci, "Emerging infectious diseases: Threats to human health and global stability," *PLoS Pathogens*, vol. 9, no. 7, e1003467, 2013.

[165] A. Morton and B. F. Finkenstädt, "Discrete time modelling of disease incidence time series by using Markov chain Monte Carlo methods," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 54, no. 3, pp. 575–594, 2005.

[166] S. D. Mylius, T. J. Hagenaars, A. K. Lugnér, and J. Wallinga, "Optimal allocation of pandemic influenza vaccine depends on age, risk and timing," *Vaccine*, vol. 26, no. 29–30, pp. 3742–3749, 2008.

[167] G. M. Nakamura, A. C. P. Monteiro, G. C. Cardoso, and A. S. Martinez, "Efficient method for comprehensive computation of agent-level epidemic dissemination in networks," *Scientific Reports*, vol. 7, p. 40885, 2017.

[168] N. Nathan, M. Barry, M. Van Herp, and H. Zeller, "Shortage of vaccines during a yellow fever outbreak in Guinea," *The Lancet*, vol. 358, no. 9299, pp. 2129–2130, 2001.

[169] J. M. Neff, J. M. Lane, V. A. Fulginiti, and D. A. Henderson, "Contact vaccinia — Transmission of vaccinia from smallpox vaccination," *The Journal of the American Medical Association*, vol. 288, no. 15, pp. 1901–1905, 2002.

[170] C. F. Nielsen, S. Kidd, A. R. Sillah, E. Davis, J. Mermin, *et al.*, "Improving burial practices and cemetery management during an Ebola virus disease epidemic — Sierra Leone, 2014," *Morbidity and Mortality Weekly Report*, vol. 64, no. 1, pp. 20–27, 2015.

[171] D. Nkoghe, M. L. Kone, A. Yada, and E. Leroy, "A limited outbreak of Ebola haemorrhagic fever in Etoumbi, Republic of Congo, 2005," *Transactions of the Royal Society of Tropical Medicine and Hygiene*, vol. 105, no. 8, pp. 466–472, 2011.

[172] E. Oehler, L. Watrin, P. Larre, I. Leparc-Goffart, S. Lastere, *et al.*, "Zika virus infection complicated by Guillain-Barre syndrome — Case report, French Polynesia, December 2013," *Eurosurveillance*, vol. 19, no. 9, p. 20720, 2014.

[173] R. S. Ostfeld, G. E. Glass, and F. Keesing, "Spatial epidemiology: An emerging (or re-emerging) discipline," *Trends in Ecology & Evolution*, vol. 20, no. 6, pp. 328–336, 2005.

[174] A. Pandey, K. E. Atkins, J. Medlock, N. Wenzel, J. P. Townsend, *et al.*, "Strategies for containing Ebola in West Africa," *Science*, vol. 346, no. 6212, pp. 991–995, 2014.

[175] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, "Epidemic processes in complex networks," *Reviews of Modern Physics*, vol. 87, no. 3, p. 925, 2015.

[176] R. Patel, I. M. Longini Jr, and M. E. Halloran, "Finding optimal vaccination strategies for pandemic influenza using genetic algorithms," *Journal of Theoretical Biology*, vol. 234, no. 2, pp. 201–212, 2005.

[177] S. Pathak, A. Maiti, and G. P. Samanta, "Rich dynamics of an SIR epidemic model," *Nonlinear Analysis: Modelling and Control*, vol. 15, no. 1, pp. 71–81, 2010.

[178] M. L. Pearson, C. B. Bridges, and S. A. Harper, "Influenza vaccination of healthcare personnel," *Morbidity and Mortality Weekly Report*, vol. 55, no. RR02, pp. 1–16, 2006.

[179] C. Poletto, M. F. C. Gomes, A. P. y Piontti, L. Rossi, L. Bioglio, *et al.*, "Assessing the impact of travel restrictions on international spread of the 2014 West African Ebola epidemic," *Eurosurveillance*, vol. 19, no. 42, p. 20936, 2014.

[180] A. Pugliese, "Population models for diseases with no recovery," *Journal of Mathematical Biology*, vol. 28, no. 1, pp. 65–82, 1990.

[181] T. D. Querec, R. S. Akondy, E. K. Lee, W. Cao, H. I. Nakaya, *et al.*, "Systems biology approach predicts immunogenicity of the yellow fever vaccine in humans," *Nature Immunology*, vol. 10, no. 1, pp. 116–125, 2009.

[182] D. M. Rao, A. Chernyakhovsky, and V. Rao, "Modeling and analysis of global epidemiology of avian influenza," *Environmental Modelling & Software*, vol. 24, no. 1, pp. 124–134, 2009.

[183] S. Riedel, "Smallpox and biological warfare: A disease revisited," *Proceedings (Baylor University. Medical Center)*, vol. 18, no. 1, pp. 13–20, 2005.

[184] S. Riley, C. Fraser, C. A. Donnelly, A. C. Ghani, L. J. Abu-Raddad, *et al.*, "Transmission dynamics of the etiological agent of SARS in Hong Kong: Impact of public health interventions," *Science*, vol. 300, no. 5627, pp. 1961–1966, 2003.

[185] C. M. Rivers, E. T. Lofgren, M. Marathe, S. Eubank, and B. L. Lewis, "Modeling the impact of interventions on an epidemic of Ebola in Sierra Leone and Liberia," *PLoS Currents Outbreaks*, Nov 6, 2014.

[186] M. G. Roberts, "A Kermack-McKendrick model applied to an infectious disease in a natural population," *Mathematical Medicine and Biology*, vol. 16, no. 4, pp. 319–332, 1999.

[187] R. Ross, *The Prevention of Malaria*. London: John Murray, 1911.

[188] M. C. Schuette and H. W. Hethcote, "Modeling the effects of varicella vaccination programs on the incidence of chickenpox and shingles," *Bulletin of Mathematical Biology*, vol. 61, no. 6, pp. 1031–1064, 1999.

[189] B. Schwartz and W. A. Orenstein, "Prioritization of pandemic influenza vaccine: Rationale and strategy for decision making," in *Vaccines for Pandemic Influenza. Current Topics in Microbiology and Immunology*, R. W. Compans and W. A. Orenstein, Eds., vol. 333, Berlin, Heidelberg: Springer, 2009, pp. 495–507.

[190] T. Sellke, "On the asymptotic distribution of the size of a stochastic epidemic," *Journal of Applied Probability*, vol. 20, no. 2, pp. 390–394, 1983.

[191] K. Shinya, M. Ebina, S. Yamada, M. Ono, N. Kasai, *et al.*, "Avian flu: Influenza virus receptors in the human airway," *Nature*, vol. 440, no. 7083, pp. 435–436, 2006.

[192] P. Shivayogi, "Vulnerable population and methods for their safeguard," *Perspectives in Clinical Research*, vol. 4, no. 1, pp. 53–57, 2013.

[193] D. I. H. Simpson, "Zika virus infection in man," *Transactions of the Royal Society of Tropical Medicine and Hygiene*, vol. 58, no. 4, pp. 335–338, 1964.

[194] T. Smith, N. Maire, A. Ross, M. Penny, N. Chitnis, *et al.*, "Towards a comprehensive simulation model of malaria epidemiology and control," *Parasitology*, vol. 135, no. 13, pp. 1507–1516, 2008.

[195] M. M. Sowilem, H. A. Kamal, and E. I. Khater, "Life table characteristics of Aedes aegypti (Diptera: Culicidae) from Saudi Arabia," *Tropical Biomedicine*, vol. 30, no. 2, pp. 301–314, 2013.

[196] J. Štěpán and D. Hlubinka, "Kermack-McKendrick epidemic model revisited," *Kybernetika*, vol. 43, no. 4, pp. 395–414, 2007.

[197] E. Tornatore, S. M. Buccellato, and P. Vetro, "Stability of a stochastic SIR system," *Physica A: Statistical Mechanics and its Applications*, vol. 354, pp. 111–126, 2005.

[198] H. C. Tuckwell and R. J. Williams, "Some properties of a simple stochastic epidemic model of SIR type," *Mathematical Biosciences*, vol. 208, no. 1, pp. 76–97, 2007.

[199] A. R. Tuite, D. N. Fisman, J. C. Kwong, and A. L. Greer, "Optimal pandemic influenza vaccine allocation strategies for the Canadian population," *PLoS One*, vol. 5, no. 5, e10520, 2010.

[200] A. R. Tuite, A. L. Greer, M. Whelan, A.-L. Winter, B. Lee, *et al.*, "Estimated epidemiologic parameters and morbidity associated with pandemic H1N1 influenza," *Canadian Medical Association Journal*, vol. 182, no. 2, pp. 131–136, 2010.

[201] J. B. Ulmer, U. Valley, and R. Rappuoli, "Vaccine manufacturing: Challenges and solutions," *Nature Biotechnology*, vol. 24, no. 11, pp. 1377–1383, 2006.

[202] L. Uscher-Pines, S. B. Omer, D. J. Barnett, T. A. Burke, and R. D. Balicer, "Priority setting for pandemic influenza: An analysis of national preparedness plans," *PLoS Medicine*, vol. 3, no. 10, e436, 2006.

[203] M. Vazeille, A. Yébakima, R. Lourenço-de Oliveira, B. Andriamahefazafy, A. Correira, *et al.*, "Oral receptivity of Aedes aegypti from Cape Verde for yellow fever, dengue, and chikungunya viruses," *Vector-Borne and Zoonotic Diseases*, vol. 13, no. 1, pp. 37–40, 2013.

[204] C. V. Ventura, M. Maia, V. Bravo-Filho, A. L. Góis, and R. Belfort Jr., "Zika virus in Brazil and macular atrophy in a child with microcephaly," *The Lancet*, vol. 387, no. 10015, p. 228, 2016.

[205] N. Vogel, C. Theisen, J. P. Leidig, J. Scripps, D. H. Graham, *et al.*, "Mining mobile datasets to enable the fine-grained stochastic simulation of Ebola diffusion," *Procedia Computer Science*, vol. 51, pp. 765–774, 2015.

[206] J. Wallinga, M. Van Boven, and M. Lipsitch, "Optimizing infectious disease interventions during an emerging epidemic," *Proceedings of the National Academy of Sciences*, vol. 107, no. 2, pp. 923–928, 2010.

[207] R. G. Webster, M. Peiris, H. Chen, and Y. Guan, "H5N1 outbreaks and enzootic influenza," *Biodiversity*, vol. 7, no. 1, pp. 51–55, 2006.

[208] L. M. Wein, D. L. Craft, and E. H. Kaplan, "Emergency response to an anthrax attack," *Proceedings of the National Academy of Sciences*, vol. 100, no. 7, pp. 4346–4351, 2003.

[209] WHO Ebola Response Team, "Ebola virus disease in West Africa — The first 9 months of the epidemic and forward projections," *The New England Journal of Medicine*, vol. 371, no. 16, pp. 1481–1495, 2014.

[210] WHO Media Center. (2018). Ebola virus disease, [Online]. Available: `http://www.who.int/mediacentre/factsheets/fs103/en/` (visited on 02/02/2018).

[211] O. P. Wójcik, J. S. Brownstein, R. Chunara, and M. A. Johansson, "Public health for the people: Participatory infectious disease surveillance in the digital age," *Emerging Themes in Epidemiology*, vol. 11, p. 7, 2014.

[212] World Health Organization, "Ebola situation reports 2015," World Health Organization, Geneva, Switzerland, Tech. Rep., 2015.

[213] ——, "Health worker Ebola infections in Guinea, Liberia and Sierra Leone," World Health Organization, Geneva, Switzerland, Tech. Rep., 2015.

[214] G. Wu and S. Yan, "Prediction of mutation trend in hemagglutinins and neuraminidases from influenza A viruses by means of cross-impact analysis," *Biochemical and Biophysical Research Communications*, vol. 326, no. 2, pp. 475–482, 2005.

[215] J. T. Wu, S. Riley, and G. M. Leung, "Spatial considerations for the allocation of pre-pandemic influenza vaccination in the United States," *Proceedings of the Royal Society of London B: Biological Sciences*, vol. 274, no. 1627, pp. 2811–2817, 2007.

[216] Z. Xu, "Traveling waves in a Kermack-McKendrick epidemic model with diffusion and latent period," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 111, pp. 66–81, 2014.

[217] G. Zaman, Y. H. Kang, and I. H. Jung, "Optimal treatment of an SIR epidemic model with time delay," *Biosystems*, vol. 98, no. 1, pp. 43–50, 2009.

[218] J. Zhou and H. W. Hethcote, "Population size dependent incidence in models for diseases without immunity," *Journal of Mathematical Biology*, vol. 32, no. 8, pp. 809–834, 1994.